(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 February 2002 (21.02.2002)

PCT

(10) International Publication Number
WO 02/14489 A2

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/US01/25054

(22) International Filing Date: 10 August 2001 (10.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/224,455 10 August 2000 (10.08.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/224,455 (CON)
Filed on 10 August 2000 (10.08.2000)

(71) Applicant *(for all designated States except US)*: **AMGEN, INC.** [US/US]; One Amgen Center Drive, Mailstop 27-4-A, Thousand Oaks, CA 91320-1789 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PASZTY, Christopher, J.** [US/US]; 1027 Westridge Drive, Ventura, CA 93003 (US). **GONG, Jianhua** [CN/US]; 2489 Gillingham Circle, Thousand Oaks, CA 91362 (US). **DAUGHERTY, Betsy** [US/US]; 3091 Stonewood Street, Simi Valley, CA 93062 (US). **ROGERS, Norma** [US/US]; 4042 Milano Place, Moorpark, CA 93021 (US).

(74) Agent: **ZUHN, Donald, L.**; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
- *without international search report and to be republished upon receipt of that report*
- -- *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/14489 A2

(54) Title: LEUCINE-RICH REPEAT-CONTAINING G-PROTEIN COUPLED RECEPTOR-8 MOLECULES AND USES THEREOF

(57) Abstract: The present invention provides Leucine-Rich Repeat-Containing G-Protein Coupled Receptor-8 (LGR8) polypeptides and nucleic acid molecules encoding the same. The invention also provides selective binding agents, vectors, host cells, and methods for producing LGR8 polypeptides. The invention further provides pharmaceutical compositions and methods for the diagnosis, treatment, amelioration, and/or prevention of diseases, disorders, and conditions associated with LGR8 polypeptides.

LEUCINE-RICH REPEAT-CONTAINING G-PROTEIN COUPLED RECEPTOR-8 MOLECULES AND USES THEREOF

This application claims the benefit of priority from U.S. Provisional Patent Application No. 60/224,455, filed on August 10, 2000, the disclosure of which is explicitly incorporated by reference herein.

Field of the Invention

The present invention relates to Leucine-Rich Repeat-Containing G-Protein Coupled Receptor-8 (LGR8) polypeptides and nucleic acid molecules encoding the same. The invention also relates to selective binding agents, vectors, host cells, and methods for producing LGR8 polypeptides. The invention further relates to pharmaceutical compositions and methods for the diagnosis, treatment, amelioration, and/or prevention of diseases, disorders, and conditions associated with LGR8 polypeptides.

Background of the Invention

Technical advances in the identification, cloning, expression, and manipulation of nucleic acid molecules and the deciphering of the human genome have greatly accelerated the discovery of novel therapeutics. Rapid nucleic acid sequencing techniques can now generate sequence information at unprecedented rates and, coupled with computational analyses, allow the assembly of overlapping sequences into partial and entire genomes and the identification of polypeptide-encoding regions. A comparison of a predicted amino acid sequence against a database compilation of known amino acid sequences allows one to determine the extent of homology to previously identified sequences and/or structural landmarks. The cloning and expression of a polypeptide-encoding region of a nucleic acid molecule provides a polypeptide product for structural and functional analyses. The manipulation of nucleic acid molecules and encoded polypeptides may confer advantageous properties on a product for use as a therapeutic.

In spite of the significant technical advances in genome research over the past decade, the potential for the development of novel therapeutics based on the human genome is still largely unrealized. Many genes encoding potentially beneficial polypeptide therapeutics or those encoding polypeptides, which may act as "targets"

The present invention further relates to four distinct LGR8 alternative splicing variants. The LGR8-A coding sequence consists of 18 coding exons that encode a large N-terminal leucine-rich repeat-containing extracellular domain, seven predicted transmembrane domains, and a cytoplasmic C-terminal region. The LGR8-B coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-B coding sequence lacks one of the exons encoding the N-terminal extracellular domain. The LGR8-C coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-C coding sequence lacks three of the exons encoding the N-terminal extracellular domain. The LGR8-D coding sequence consists of exons encoding approximately 90% of the N-terminal extracellular domain of the LGR8-B coding sequence, but lacks exons encoding the transmembrane domains and the cytoplasmic C-terminal region. Thus, LGR8-D is a secreted, N-terminal extracellular domain version of LGR8-B and likely functions as an antagonist of the LGR8 signaling pathway. LGR8-D is truncated very near the C-terminal end of the N-terminal extracellular domain by virtue of the fact that an additional exon, which contains stop codons, is spliced in just 5' of the exon which encodes the first transmembrane domain of LGR8-A, LGR8-B, and LGR8-C. It is likely that the N-terminal extracellular domains of LGR8-A, LGR8-B, and LGR8-C would be able to function as antagonists of the LGR8 signaling pathway.

The invention provides for an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22;

(b) a nucleotide sequence encoding the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(c) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of either (a) or (b); and

(d) a nucleotide sequence complementary to either (a) or (b).

(a) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one conservative amino acid substitution, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(b) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid insertion, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(c) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid deletion, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(d) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 which has a C- and/or N- terminal truncation, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ

(c) an amino acid sequence which is at least about 70 percent identical to the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(d) a fragment of the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 comprising at least about 25 amino acid residues, wherein the fragment has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or is antigenic; and

(e) an amino acid sequence for an allelic variant or splice variant of the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or any of (a) - (c).

The invention further provides for an isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

(a) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

The invention still further provides for an isolated polypeptide comprising the amino acid sequence as set forth in SEQ ID NO: 2 with at least one conservative amino acid substitution selected from the group consisting of: isoleucine at position 26; valine at position 41; isoleucine at position 55; aspartic acid at position 78; aspartic acid at position 123; arginine at position 130; valine at position 135; methionine at position 142; leucine at position 166; tyrosine at position 167; lysine at position 201; valine at position 204; isoleucine at position 216; glutamatic acid at position 217; leucine at position 221; leucine at position 240; leucine at position 252; isoleucine at position 277; methionine at position 288; lysine at position 290; isoleucine at position 324; isoleucine at position 341; isoleucine at position 344; aspartic acid at position 350; leucine at position 376; valine at position 420; valine at position 425; valine at position 427; isoleucine at position 434; tyrosine at position 442; arginine at position 444; tyrosine at position 450; isoleucine at position 466; isoleucine at position 471; leucine at position 476; phenylalanine at position 478; glutamatic acid at position 481; histidine at position 485; phenylalanine at position 515; tyrosine at position 521; isoleucine at position 522; tyrosine at position 526; valine at position 531; valine at position 541; isoleucine at position 551; valine at position 552; glutamatic acid at position 561; phenylalanine at position 562; tyrosine at position 566; tyrosine at position 577; aspartic acid at position 579; isoleucine at position 597; isoleucine at position 603; valine at position 616; isoleucine at position 621; isoleucine at position 626; lysine at position 632; leucine at position 649; isoleucine at position 654; valine at position 675; isoleucine at position 682; glutamatic acid at position 700; isoleucine at position 702; tyrosine at position 707; tyrosine at position 709; isoleucine at position 727; valine at position 729; methionine at position 737; methionine at position 745; and leucine at position 749; wherein the polypeptide has an activity of the polypeptide as set forth in SEQ ID NO: 2.

Also provided are fusion polypeptides comprising LGR8 amino acid sequences.

The present invention also provides for an expression vector comprising the isolated nucleic acid molecules as set forth herein, recombinant host cells comprising the recombinant nucleic acid molecules as set forth herein, and a method of producing

Methods of regulating expression and modulating (*i.e.*, increasing or decreasing) levels of an LGR8 polypeptide are also encompassed by the invention. One method comprises administering to an animal a nucleic acid molecule encoding an LGR8 polypeptide. In another method, a nucleic acid molecule comprising elements that regulate or modulate the expression of an LGR8 polypeptide may be administered. Examples of these methods include gene therapy, cell therapy, and anti-sense therapy as further described herein.

LGR8 polypeptides can be used for identifying ligands thereof. Various forms of "expression cloning" have been used for cloning ligands for receptors (*See, e.g.*, Davis *et al.*, 1996, *Cell*, 87:1161-69). These and other LGR8 ligand cloning experiments are described in greater detail herein. Isolation of the LGR8 ligand(s) allows for the identification or development of novel agonists and/or antagonists of the LGR8 signaling pathway. Such agonists and antagonists include LGR8 ligand(s), anti-LGR8 ligand antibodies and derivatives thereof, small molecules, or antisense oligonucleotides, any of which can be used for potentially treating one or more diseases or disorders, including those recited herein.

Brief Description of the Figures

Figures 1A-1D illustrate a nucleotide sequence (SEQ ID NO: 1) encoding human LGR8-A (SEQ ID NO: 2). The predicted signal sequence is indicated (underline);

Figures 2A-2B illustrate a nucleotide sequence (SEQ ID NO: 4) encoding the N-terminal extracellular domain (absent the signal peptide) of human LGR8-A (SEQ ID NO: 5);

Figures 3A-3D illustrate a nucleotide sequence (SEQ ID NO: 6) encoding human LGR8-B (SEQ ID NO: 7). The predicted signal sequence is indicated (underline);

Figures 4A-4B illustrate a nucleotide sequence (SEQ ID NO: 9) encoding the N-terminal extracellular domain (absent the signal peptide) of human LGR8-B (SEQ ID NO: 10);

Figures 5A-5D illustrate a nucleotide sequence (SEQ ID NO: 11) encoding human LGR8-C (SEQ ID NO: 12). The predicted signal sequence is indicated (underline);

transmembrane domains, and a C-terminal cytoplasmic domain. Accordingly, LGR8-A, LGR8-B, and LGR8-C are useful as targets for agonistic or antagonistic molecules, including, but not limited to, antibodies, fusion polypeptides, carbohydrates, polynucleotides (such as antisense oligonucleotides), or small molecular weight organic molecules.

Additionally, it will be understood that the N-terminal extracellular domains of LGR8-A, LGR8-B, and LGR8-C can be used as antagonists of the LGR8 signaling pathway, for example, where the N-terminal extracellular domain is fused to an Fc portion of an antibody.

It will also be appreciated that LGR8-D is a secreted form of the N-terminal extracellular domain of LGR8-B. In this regard LGR8-D may act as an antagonist of the LGR8-B ligand(s). LGR8-D can be used as a target for antagonistic and agonistic molecules, including, but not limited to, antibodies, fusion polypeptides, carbohydrates, polynucleotides (such as antisense oligonucleotides), or small molecular weight organic molecules. For example, an antagonist specific for LGR8-D would inhibit the antagonistic activity of LGR8-D, thus enhancing the activity of LGR8-D ligand(s) and/or enhancing signaling through LGR8 receptors. Conversely an agonist specific for LGR8-D would enhance the antagonistic activity of LGR8-D, thus diminishing the activity of LGR8-D ligand(s) and/or diminishing signaling through LGR8.

Definitions

The terms "LGR8 gene" or "LGR8 nucleic acid molecule" or "LGR8 polynucleotide" refer to a nucleic acid molecule comprising or consisting of a nucleotide sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, a nucleotide sequence encoding the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, and nucleic acid molecules as defined herein.

The term "LGR8 polypeptide allelic variant" refers to one of several possible naturally occurring alternate forms of a gene occupying a given locus on a chromosome of an organism or a population of organisms.

The term "vector" is used to refer to any molecule (*e.g.*, nucleic acid, plasmid, or virus) used to transfer coding information to a host cell.

The term "expression vector" refers to a vector that is suitable for transformation of a host cell and contains nucleic acid sequences that direct and/or control the expression of inserted heterologous nucleic acid sequences. Expression includes, but is not limited to, processes such as transcription, translation, and RNA splicing, if introns are present.

The term "operably linked" is used herein to refer to an arrangement of flanking sequences wherein the flanking sequences so described are configured or assembled so as to perform their usual function. Thus, a flanking sequence operably linked to a coding sequence may be capable of effecting the replication, transcription and/or translation of the coding sequence. For example, a coding sequence is operably linked to a promoter when the promoter is capable of directing transcription of that coding sequence. A flanking sequence need not be contiguous with the coding sequence, so long as it functions correctly. Thus, for example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

The term "host cell" is used to refer to a cell which has been transformed, or is capable of being transformed with a nucleic acid sequence and then of expressing a selected gene of interest. The term includes the progeny of the parent cell, whether or not the progeny is identical in morphology or in genetic make-up to the original parent, so long as the selected gene is present.

The term "LGR8 polypeptide" refers to a polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 and related polypeptides. Related polypeptides include LGR8 polypeptide fragments, LGR8 polypeptide orthologs, LGR8 polypeptide variants, and LGR8 polypeptide derivatives, which possess at least one activity of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. LGR8 polypeptides may be mature polypeptides, as defined herein, and may or may

polypeptides) as compared to the LGR8 polypeptide amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 (with or without a leader sequence). Variants may be naturally occurring (*e.g.*, LGR8 polypeptide allelic variants, LGR8 polypeptide orthologs, and LGR8 polypeptide splice variants) or artificially constructed. Such LGR8 polypeptide variants may be prepared from the corresponding nucleic acid molecules having a DNA sequence that varies accordingly from the DNA sequence as set forth in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22. In preferred embodiments, the variants have from 1 to 3, or from 1 to 5, or from 1 to 10, or from 1 to 15, or from 1 to 20, or from 1 to 25, or from 1 to 50, or from 1 to 75, or from 1 to 100, or more than 100 amino acid substitutions, insertions, additions and/or deletions, wherein the substitutions may be conservative, or non-conservative, or any combination thereof.

The term "LGR8 polypeptide derivatives" refers to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, LGR8 polypeptide fragments, LGR8 polypeptide orthologs, or LGR8 polypeptide variants, as defined herein, that have been chemically modified. The term "LGR8 polypeptide derivatives" also refers to the polypeptides encoded by LGR8 polypeptide allelic variants or LGR8 polypeptide splice variants, as defined herein, that have been chemically modified.

The term "mature LGR8 polypeptide" refers to an LGR8 polypeptide lacking a leader sequence. A mature LGR8 polypeptide may also include other modifications such as proteolytic processing of the amino-terminus (with or without a leader sequence) and/or the carboxyl-terminus, cleavage of a smaller polypeptide from a larger precursor, N-linked and/or O-linked glycosylation, and the like. Exemplary mature CHL2 polypeptides are depicted by the amino acid sequences of SEQ ID NO: 3, SEQ ID NO: 8, SEQ ID NO: 13, SEQ ID NO: 18, OR SEQ ID NO: 21.

The term "LGR8 fusion polypeptide" refers to a fusion of one or more amino acids (such as a heterologous protein or peptide) at the amino- or carboxyl-terminus of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO:

alignments (if any) addressed by a particular mathematical model or computer program (*i.e.*, "algorithms").

The term "similarity" is a related concept, but in contrast to "identity," "similarity" refers to a measure of relatedness that includes both identical matches and conservative substitution matches. If two polypeptide sequences have, for example, 10/20 identical amino acids, and the remainder are all non-conservative substitutions, then the percent identity and similarity would both be 50%. If in the same example, there are five more positions where there are conservative substitutions, then the percent identity remains 50%, but the percent similarity would be 75% (15/20). Therefore, in cases where there are conservative substitutions, the percent similarity between two polypeptides will be higher than the percent identity between those two polypeptides.

The term "naturally occurring" or "native" when used in connection with biological materials such as nucleic acid molecules, polypeptides, host cells, and the like, refers to materials which are found in nature and are not manipulated by man. Similarly, "non-naturally occurring" or "non-native" as used herein refers to a material that is not found in nature or that has been structurally modified or synthesized by man.

The terms "effective amount" and "therapeutically effective amount" each refer to the amount of an LGR8 polypeptide or LGR8 nucleic acid molecule used to support an observable level of one or more biological activities of the LGR8 polypeptides as set forth herein.

The term "pharmaceutically acceptable carrier" or "physiologically acceptable carrier" as used herein refers to one or more formulation materials suitable for accomplishing or enhancing the delivery of the LGR8 polypeptide, LGR8 nucleic acid molecule, or LGR8 selective binding agent as a pharmaceutical composition.

The term "antigen" refers to a molecule or a portion of a molecule capable of being bound by a selective binding agent, such as an antibody, and additionally capable of being used in an animal to produce antibodies capable of binding to an epitope of that antigen. An antigen may have one or more epitopes.

The term "selective binding agent" refers to a molecule or molecules having specificity for an LGR8 polypeptide. As used herein, the terms, "specific" and "specificity" refer to the ability of the selective binding agents to bind to human LGR8 polypeptides and not to bind to human non-LGR8 polypeptides. It will be

residues compared to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. Such related LGR8 polypeptides may comprise, for example, an addition and/or a deletion of one or more N-linked or O-linked glycosylation sites or an addition and/or a deletion of one or more cysteine residues.

Related nucleic acid molecules also include fragments of LGR8 nucleic acid molecules which encode a polypeptide of at least about 25 contiguous amino acids, or about 50 amino acids, or about 75 amino acids, or about 100 amino acids, or about 150 amino acids, or about 200 amino acids, or more than 200 amino acid residues of the LGR8 polypeptide of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

In addition, related LGR8 nucleic acid molecules also include those molecules which comprise nucleotide sequences which hybridize under moderately or highly stringent conditions as defined herein with the fully complementary sequence of the LGR8 nucleic acid molecule of any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or of a molecule encoding a polypeptide, which polypeptide comprises the amino acid sequence as shown in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or of a nucleic acid fragment as defined herein, or of a nucleic acid fragment encoding a polypeptide as defined herein. Hybridization probes may be prepared using the LGR8 sequences provided herein to screen cDNA, genomic or synthetic DNA libraries for related sequences. Regions of the DNA and/or amino acid sequence of LGR8 polypeptide that exhibit significant identity to known sequences are readily determined using sequence alignment algorithms as described herein and those regions may be used to design probes for screening.

The term "highly stringent conditions" refers to those conditions that are designed to permit hybridization of DNA strands whose sequences are highly complementary, and to exclude hybridization of significantly mismatched DNAs.

stringent conditions" are 0.015 M sodium chloride, 0.0015 M sodium citrate at 50-65°C or 0.015 M sodium chloride, 0.0015 M sodium citrate, and 20% formamide at 37-50°C. By way of example, "moderately stringent conditions" of 50°C in 0.015 M sodium ion will allow about a 21% mismatch.

It will be appreciated by those skilled in the art that there is no absolute distinction between "highly stringent conditions" and "moderately stringent conditions." For example, at 0.015 M sodium ion (no formamide), the melting temperature of perfectly matched long DNA is about 71°C. With a wash at 65°C (at the same ionic strength), this would allow for approximately a 6% mismatch. To capture more distantly related sequences, one skilled in the art can simply lower the temperature or raise the ionic strength.

A good estimate of the melting temperature in 1M NaCl* for oligonucleotide probes up to about 20nt is given by:

Tm = 2°C per A-T base pair + 4°C per G-C base pair

*The sodium ion concentration in 6X salt sodium citrate (SSC) is 1M. *See* Suggs *et al.*, *Developmental Biology Using Purified Genes* 683 (Brown and Fox, eds., 1981).

High stringency washing conditions for oligonucleotides are usually at a temperature of 0-5°C below the Tm of the oligonucleotide in 6X SSC, 0.1% SDS.

In another embodiment, related nucleic acid molecules comprise or consist of a nucleotide sequence that is at least about 70 percent identical to the nucleotide sequence as shown in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or comprise or consist essentially of a nucleotide sequence encoding a polypeptide that is at least about 70 percent identical to the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. In preferred embodiments, the nucleotide sequences are about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the nucleotide sequence as shown in any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or the nucleotide sequences encode a polypeptide that is about 75 percent, or about 80 percent, or about 85 percent, or about 90 percent, or about 95, 96, 97, 98, or 99 percent identical to the polypeptide sequence as set forth in

Conservative amino acid substitutions also encompass non-naturally occurring amino acid residues that are typically incorporated by chemical peptide synthesis rather than by synthesis in biological systems. These include peptidomimetics, and other reversed or inverted forms of amino acid moieties.

Naturally occurring residues may be divided into classes based on common side chain properties:

1) hydrophobic: norleucine, Met, Ala, Val, Leu, Ile;

2) neutral hydrophilic: Cys, Ser, Thr;

3) acidic: Asp, Glu;

4) basic: Asn, Gln, His, Lys, Arg;

5) residues that influence chain orientation: Gly, Pro; and

6) aromatic: Trp, Tyr, Phe.

For example, non-conservative substitutions may involve the exchange of a member of one of these classes for a member from another class. Such substituted residues may be introduced into regions of the human LGR8 polypeptide that are homologous with non-human LGR8 polypeptides, or into the non-homologous regions of the molecule.

In making such changes, the hydropathic index of amino acids may be considered. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. The hydropathic indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte *et al.*, 1982, *J. Mol. Biol.* 157:105-31). It is known that certain amino acids may be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity, particularly where the biologically

Gly	Pro, Ala	Ala
His	Asn, Gln, Lys, Arg	Arg
Ile	Leu, Val, Met, Ala, Phe, Norleucine	Leu
Leu	Norleucine, Ile, Val, Met, Ala, Phe	Ile
Lys	Arg, 1,4 Diamino-butyric Acid, Gln, Asn	Arg
Met	Leu, Phe, Ile	Leu
Phe	Leu, Val, Ile, Ala, Tyr	Leu
Pro	Ala	Gly
Ser	Thr, Ala, Cys	Thr
Thr	Ser	Ser
Trp	Tyr, Phe	Tyr
Tyr	Trp, Phe, Thr, Ser	Phe
Val	Ile, Met, Leu, Phe, Ala, Norleucine	Leu

A skilled artisan will be able to determine suitable variants of the polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 using well-known techniques. For identifying suitable areas of the molecule that may be changed without destroying biological activity, one skilled in the art may target areas not believed to be important for activity. For example, when similar polypeptides with similar activities from the same species or from other species are known, one skilled in the art may compare the amino acid sequence of an LGR8 polypeptide to such similar polypeptides. With such a comparison, one can identify residues and portions of the molecules that are conserved among similar polypeptides. It will be appreciated that changes in areas of the LGR8 molecule that are not conserved relative to such similar polypeptides would be less likely to adversely affect the biological activity and/or structure of an LGR8 polypeptide. One skilled in

computer programs are currently available to assist with predicting secondary structure. One method of predicting secondary structure is based upon homology modeling. For example, two polypeptides or proteins that have a sequence identity of greater than 30%, or similarity greater than 40%, often have similar structural topologies. The recent growth of the protein structural database (PDB) has provided enhanced predictability of secondary structure, including the potential number of folds within the structure of a polypeptide or protein. *See* Holm *et al.*, 1999, *Nucleic Acids Res.* 27:244-47. It has been suggested that there are a limited number of folds in a given polypeptide or protein and that once a critical number of structures have been resolved, structural prediction will become dramatically more accurate (Brenner *et al.*, 1997, *Curr. Opin. Struct. Biol.* 7:369-76).

Additional methods of predicting secondary structure include "threading" (Jones, 1997, *Curr. Opin. Struct. Biol.* 7:377-87; Sippl *et al.*, 1996, *Structure* 4:15-19), "profile analysis" (Bowie *et al.*, 1991, *Science*, 253:164-70; Gribskov *et al.*, 1990, *Methods Enzymol.* 183:146-59; Gribskov *et al.*, 1987, *Proc. Nat. Acad. Sci. U.S.A.* 84:4355-58), and "evolutionary linkage" (*See* Holm *et al.*, *supra*, and Brenner *et al.*, *supra*).

Preferred LGR8 polypeptide variants include glycosylation variants wherein the number and/or type of glycosylation sites have been altered compared to the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. In one embodiment, LGR8 polypeptide variants comprise a greater or a lesser number of N-linked glycosylation sites than the amino acid sequence set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. An N-linked glycosylation site is characterized by the sequence: Asn-X-Ser or Asn-X-Thr, wherein the amino acid residue designated as X may be any amino acid residue except proline. The substitution of amino acid residues to create this sequence provides a potential new site for the addition of an N-linked carbohydrate chain. Alternatively, substitutions that eliminate this sequence will remove an existing N-linked carbohydrate chain. Also provided is a rearrangement of N-linked carbohydrate chains wherein one or more N-linked glycosylation sites (typically those

SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23. Related nucleic acid molecules also comprise or consist of a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one modification selected from the group consisting of amino acid substitutions, amino acid insertions, amino acid deletions, carboxyl-terminal truncations, and amino-terminal truncations and wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

In addition, the polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide, may be fused to a homologous polypeptide to form a homodimer or to a *heterologous polypeptide to form a heterodimer.* *Heterologous peptides and* polypeptides include, but are not limited to: an epitope to allow for the detection and/or isolation of an LGR8 fusion polypeptide; a transmembrane receptor protein or a portion thereof, such as an extracellular domain or a transmembrane and intracellular domain; a ligand or a portion thereof which binds to a transmembrane receptor protein; an enzyme or portion thereof which is catalytically active; a polypeptide or peptide which promotes oligomerization, such as a leucine zipper domain; a polypeptide or peptide which increases stability, such as an immunoglobulin constant region; and a polypeptide which has a therapeutic activity different from the polypeptide comprising the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide.

IgG1	TNF receptor	septic shock	Fisher *et al.*, 1996, *N. Engl. J. Med.* 334:1697-1702; Van Zee *et al.*, 1996, *J. Immunol.* 156:2221-30
IgG, IgA, IgM, or IgE (excluding the first domain)	TNF receptor	inflammation, autoimmune disorders	U.S. Patent No. 5,808,029
IgG1	CD4 receptor	AIDS	Capon *et al.*, 1989, *Nature* 337: 525-31
IgG1, IgG3	N-terminus of IL-2	anti-cancer, antiviral	Harvill *et al.*, 1995, *Immunotech.* 1:95-105
IgG1	C-terminus of OPG	osteoarthritis; bone density	WO 97/23614
IgG1	N-terminus of leptin	anti-obesity	PCT/US 97/23183, filed December 11, 1997
Human Ig Cγ1	CTLA-4	autoimmune disorders	Linsley, 1991, *J. Exp. Med.*, 174:561-69

In one example, a human IgG hinge, CH2, and CH3 region may be fused at either the amino-terminus or carboxyl-terminus of the LGR8 polypeptides using methods known to the skilled artisan. In another example, a human IgG hinge, CH2, and CH3 region may be fused at either the amino-terminus or carboxyl-terminus of an LGR8 polypeptide fragment (*e.g.*, the predicted extracellular portion of LGR8 polypeptide).

The resulting LGR8 fusion polypeptide may be purified by use of a Protein A affinity column. Peptides and proteins fused to an Fc region have been found to exhibit a substantially greater half-life *in vivo* than the unfused counterpart. Also, a fusion to an Fc region allows for dimerization/multimerization of the fusion polypeptide. The Fc region may be a naturally occurring Fc region, or may be altered to improve certain qualities, such as therapeutic qualities, circulation time, or reduced aggregation.

Identity and similarity of related nucleic acid molecules and polypeptides are readily calculated by known methods. Such methods include, but are not limited to those described in *Computational Molecular Biology* (A.M. Lesk, ed., Oxford University Press 1988); *Biocomputing: Informatics and Genome Projects* (D.W. Smith, ed., Academic Press 1993); *Computer Analysis of Sequence Data* (Part 1, A.M. Griffin and H.G. Griffin, eds., Humana Press 1994); G. von Heinle, *Sequence Analysis in Molecular Biology* (Academic Press 1987); *Sequence Analysis Primer* (M.

Algorithm: Needleman and Wunsch, 1970, *J. Mol. Biol.* 48:443-53;
Comparison matrix: BLOSUM 62 (Henikoff *et al.*, *supra*);
Gap Penalty: 12
Gap Length Penalty: 4
Threshold of Similarity: 0

The GAP program is useful with the above parameters. The aforementioned parameters are the default parameters for polypeptide comparisons (along with no penalty for end gaps) using the GAP algorithm.

Preferred parameters for nucleic acid molecule sequence comparison include the following:

Algorithm: Needleman and Wunsch, *supra*;
Comparison matrix: matches = +10, mismatch = 0
Gap Penalty: 50
Gap Length Penalty: 3

The GAP program is also useful with the above parameters. The aforementioned parameters are the default parameters for nucleic acid molecule comparisons.

Other exemplary algorithms, gap opening penalties, gap extension penalties, comparison matrices, and thresholds of similarity may be used, including those set forth in the Program Manual, Wisconsin Package, Version 9, September, 1997. The particular choices to be made will be apparent to those of skill in the art and will depend on the specific comparison to be made, such as DNA-to-DNA, protein-to-protein, protein-to-DNA; and additionally, whether the comparison is between given pairs of sequences (in which case GAP or BestFit are generally preferred) or between one sequence and a large database of sequences (in which case FASTA or BLASTA are preferred).

Nucleic Acid Molecules

The nucleic acid molecules encoding a polypeptide comprising the amino acid sequence of an LGR8 polypeptide can readily be obtained in a variety of ways

introducing the expression vector into an appropriate host, the encoded LGR8 polypeptide may be produced in large amounts.

Another method for obtaining a suitable nucleic acid sequence is the polymerase chain reaction (PCR). In this method, cDNA is prepared from poly(A)+RNA or total RNA using the enzyme reverse transcriptase. Two primers, typically complementary to two separate regions of cDNA encoding the amino acid sequence of an LGR8 polypeptide, are then added to the cDNA along with a polymerase such as *Taq* polymerase, and the polymerase amplifies the cDNA region between the two primers.

Another means of preparing a nucleic acid molecule encoding the amino acid sequence of an LGR8 polypeptide is chemical synthesis using methods well known to the skilled artisan such as those described by Engels *et al.*, 1989, *Angew. Chem. Intl. Ed.* 28:716-34. These methods include, *inter alia*, the phosphotriester, phosphoramidite, and H-phosphonate methods for nucleic acid synthesis. A preferred method for such chemical synthesis is polymer-supported synthesis using standard phosphoramidite chemistry. Typically, the DNA encoding the amino acid sequence of an LGR8 polypeptide will be several hundred nucleotides in length. Nucleic acids larger than about 100 nucleotides can be synthesized as several fragments using these methods. The fragments can then be ligated together to form the full-length nucleotide sequence of an LGR8 gene. Usually, the DNA fragment encoding the amino-terminus of the polypeptide will have an ATG, which encodes a methionine residue. This methionine may or may not be present on the mature form of the LGR8 polypeptide, depending on whether the polypeptide produced in the host cell is designed to be secreted from that cell. Other methods known to the skilled artisan may be used as well.

In certain embodiments, nucleic acid variants contain codons which have been altered for optimal expression of an LGR8 polypeptide in a given host cell. Particular codon alterations will depend upon the LGR8 polypeptide and host cell selected for expression. Such "codon optimization" can be carried out by a variety of methods, for example, by selecting codons which are preferred for use in highly expressed genes in a given host cell. Computer algorithms which incorporate codon frequency tables such as "Eco_high.Cod" for codon preference of highly expressed bacterial genes may be used and are provided by the University of Wisconsin Package Version 9.0 (Genetics Computer Group, Madison, WI). Other useful codon frequency tables

to be expressed, and a selectable marker element. Each of these sequences is discussed below.

Optionally, the vector may contain a "tag"-encoding sequence, *i.e.*, an oligonucleotide molecule located at the 5' or 3' end of the LGR8 polypeptide coding sequence; the oligonucleotide sequence encodes polyHis (such as hexaHis), or another "tag" such as FLAG, HA (hemaglutinin influenza virus), or *myc* for which commercially available antibodies exist. This tag is typically fused to the polypeptide upon expression of the polypeptide, and can serve as a means for affinity purification of the LGR8 polypeptide from the host cell. Affinity purification can be accomplished, for example, by column chromatography using antibodies against the tag as an affinity matrix. Optionally, the tag can subsequently be removed from the purified LGR8 polypeptide by various means such as using certain peptidases for cleavage.

Flanking sequences may be homologous (*i.e.*, from the same species and/or strain as the host cell), heterologous (*i.e.*, from a species other than the host cell species or strain), hybrid (*i.e.*, a combination of flanking sequences from more than one source), or synthetic, or the flanking sequences may be native sequences that normally function to regulate LGR8 polypeptide expression. As such, the source of a flanking sequence may be any prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the flanking sequence is functional in, and can be activated by, the host cell machinery.

Flanking sequences useful in the vectors of this invention may be obtained by any of several methods well known in the art. Typically, flanking sequences useful herein – other than the LGR8 gene flanking sequences – will have been previously identified by mapping and/or by restriction endonuclease digestion and can thus be isolated from the proper tissue source using the appropriate restriction endonucleases. In some cases, the full nucleotide sequence of a flanking sequence may be known. Here, the flanking sequence may be synthesized using the methods described herein for nucleic acid synthesis or cloning.

Where all or only a portion of the flanking sequence is known, it may be obtained using PCR and/or by screening a genomic library with a suitable oligonucleotide and/or flanking sequence fragment from the same or another species. Where the flanking sequence is not known, a fragment of DNA containing a flanking sequence may be isolated from a larger piece of DNA that may contain, for example,

Other selection genes may be used to amplify the gene that will be expressed. Amplification is the process wherein genes that are in greater demand for the production of a protein critical for growth are reiterated in tandem within the chromosomes of successive generations of recombinant cells. Examples of suitable selectable markers for mammalian cells include dihydrofolate reductase (DHFR) and thymidine kinase. The mammalian cell transformants are placed under selection pressure wherein only the transformants are uniquely adapted to survive by virtue of the selection gene present in the vector. Selection pressure is imposed by culturing the transformed cells under conditions in which the concentration of selection agent in the medium is successively changed, thereby leading to the amplification of both the selection gene and the DNA that encodes an LGR8 polypeptide. As a result, increased quantities of LGR8 polypeptide are synthesized from the amplified DNA.

A ribosome binding site is usually necessary for translation initiation of mRNA and is characterized by a Shine-Dalgarno sequence (prokaryotes) or a Kozak sequence (eukaryotes). The element is typically located 3' to the promoter and 5' to the coding sequence of an LGR8 polypeptide to be expressed. The Shine-Dalgarno sequence is varied but is typically a polypurine (*i.e.*, having a high A-G content). Many Shine-Dalgarno sequences have been identified, each of which can be readily synthesized using methods set forth herein and used in a prokaryotic vector.

A leader, or signal, sequence may be used to direct an LGR8 polypeptide out of the host cell. Typically, a nucleotide sequence encoding the signal sequence is positioned in the coding region of an LGR8 nucleic acid molecule, or directly at the 5' end of an LGR8 polypeptide coding region. Many signal sequences have been identified, and any of those that are functional in the selected host cell may be used in conjunction with an LGR8 nucleic acid molecule. Therefore, a signal sequence may be homologous (naturally occurring) or heterologous to the LGR8 nucleic acid molecule. Additionally, a signal sequence may be chemically synthesized using methods described herein. In most cases, the secretion of an LGR8 polypeptide from the host cell via the presence of a signal peptide will result in the removal of the signal peptide from the secreted LGR8 polypeptide. The signal sequence may be a component of the vector, or it may be a part of an LGR8 nucleic acid molecule that is inserted into the vector.

Included within the scope of this invention is the use of either a nucleotide sequence encoding a native LGR8 polypeptide signal sequence joined to an LGR8

not interrupt the coding sequence. Any intron from any source, including viral, prokaryotic and eukaryotic (plant or animal) organisms, may be used to practice this invention, provided that it is compatible with the host cell into which it is inserted. Also included herein are synthetic introns. Optionally, more than one intron may be used in the vector.

The expression and cloning vectors of the present invention will typically contain a promoter that is recognized by the host organism and operably linked to the molecule encoding the LGR8 polypeptide. Promoters are untranscribed sequences located upstream (*i.e.*, 5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control the transcription of the structural gene. Promoters are conventionally grouped into one of two classes: inducible promoters and constitutive promoters. Inducible promoters initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, such as the presence or absence of a nutrient or a change in temperature. Constitutive promoters, on the other hand, initiate continual gene product production; that is, there is little or no control over gene expression. A large number of promoters, recognized by a variety of potential host cells, are well known. A suitable promoter is operably linked to the DNA encoding LGR8 polypeptide by removing the promoter from the source DNA by restriction enzyme digestion and inserting the desired promoter sequence into the vector. The native LGR8 promoter sequence may be used to direct amplification and/or expression of an LGR8 nucleic acid molecule. A heterologous promoter is preferred, however, if it permits greater transcription and higher yields of the expressed protein as compared to the native promoter, and if it is compatible with the host cell system that has been selected for use.

Promoters suitable for use with prokaryotic hosts include the beta-lactamase and lactose promoter systems; alkaline phosphatase; a tryptophan (trp) promoter system; and hybrid promoters such as the tac promoter. Other known bacterial promoters are also suitable. Their sequences have been published, thereby enabling one skilled in the art to ligate them to the desired DNA sequence, using linkers or adapters as needed to supply any useful restriction sites.

Suitable promoters for use with yeast hosts are also well known in the art. Yeast enhancers are advantageously used with yeast promoters. Suitable promoters for use with mammalian host cells are well known and include, but are not limited to, those obtained from the genomes of viruses such as polyoma virus, fowlpox virus,

gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason *et al.*, 1986, *Science* 234:1372-78).

An enhancer sequence may be inserted into the vector to increase the transcription of a DNA encoding an LGR8 polypeptide of the present invention by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about 10-300 bp in length, that act on the promoter to increase transcription. Enhancers are relatively orientation and position independent. They have been found 5' and 3' to the transcription unit. Several enhancer sequences available from mammalian genes are known (*e.g.*, globin, elastase, albumin, alpha-feto-protein and insulin). Typically, however, an enhancer from a virus will be used. The SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer, and adenovirus enhancers are exemplary enhancing elements for the activation of eukaryotic promoters. While an enhancer may be spliced into the vector at a position 5' or 3' to an LGR8 nucleic acid molecule, it is typically located at a site 5' from the promoter.

Expression vectors of the invention may be constructed from a starting vector such as a commercially available vector. Such vectors may or may not contain all of the desired flanking sequences. Where one or more of the flanking sequences described herein are not already present in the vector, they may be individually obtained and ligated into the vector. Methods used for obtaining each of the flanking sequences are well known to one skilled in the art.

Preferred vectors for practicing this invention are those that are compatible with bacterial, insect, and mammalian host cells. Such vectors include, *inter alia*, pCRII, pCR3, and pcDNA3.1 (Invitrogen, San Diego, CA), pBSII (Stratagene, La Jolla, CA), pET15 (Novagen, Madison, WI), pGEX (Pharmacia Biotech, Piscataway, NJ), pEGFP-N2 (Clontech, Palo Alto, CA), pETL (BlueBacII, Invitrogen), pDSR-alpha (PCT Pub. No. WO 90/14363) and pFastBacDual (Gibco-BRL, Grand Island, NY).

Additional suitable vectors include, but are not limited to, cosmids, plasmids, or modified viruses, but it will be appreciated that the vector system must be compatible with the selected host cell. Such vectors include, but are not limited to plasmids such as Bluescript® plasmid derivatives (a high copy number ColE1-based phagemid; Stratagene Cloning Systems, La Jolla CA), PCR cloning plasmids designed for cloning Taq-amplified PCR products (*e.g.*, TOPO™ TA Cloning® Kit

may contain a dominantly acting selection gene. Other suitable mammalian cell lines include but are not limited to, mouse neuroblastoma N2A cells, HeLa, mouse L-929 cells, 3T3 lines derived from Swiss, Balb-c or NIH mice, BHK or HaK hamster cell lines. Each of these cell lines is known by and available to those skilled in the art of protein expression.

Similarly useful as host cells suitable for the present invention are bacterial cells. For example, the various strains of *E. coli* (*e.g.*, HB101, DH5α, DH10, and MC1061) are well-known as host cells in the field of biotechnology. Various strains of *B. subtilis*, *Pseudomonas spp.*, other *Bacillus spp.*, *Streptomyces spp.*, and the like may also be employed in this method.

Many strains of yeast cells known to those skilled in the art are also available as host cells for the expression of the polypeptides of the present invention. Preferred yeast cells include, for example, *Saccharomyces cerivisae* and *Pichia pastoris*.

Additionally, where desired, insect cell systems may be utilized in the methods of the present invention. Such systems are described, for example, in Kitts *et al.*, 1993, *Biotechniques*, 14:810-17; Lucklow, 1993, *Curr. Opin. Biotechnol.* 4:564-72; and Lucklow *et al.*, 1993, *J. Virol.*, 67:4566-79. Preferred insect cells are Sf-9 and Hi5 (Invitrogen).

One may also use transgenic animals to express glycosylated LGR8 polypeptides. For example, one may use a transgenic milk-producing animal (a cow or goat, for example) and obtain the present glycosylated polypeptide in the animal milk. One may also use plants to produce LGR8 polypeptides, however, in general, the glycosylation occurring in plants is different from that produced in mammalian cells, and may result in a glycosylated product which is not suitable for human therapeutic use.

Polypeptide Production

Host cells comprising an LGR8 polypeptide expression vector may be cultured using standard media well known to the skilled artisan. The media will usually contain all nutrients necessary for the growth and survival of the cells. Suitable media for culturing *E. coli* cells include, for example, Luria Broth (LB) and/or Terrific Broth (TB). Suitable media for culturing eukaryotic cells include Roswell Park Memorial Institute medium 1640 (RPMI 1640), Minimal Essential Medium (MEM) and/or Dulbecco's Modified Eagle Medium (DMEM), all of which

LGR8 polypeptide can then be analyzed using gel electrophoresis, immunoprecipitation, or the like. If it is desired to isolate the LGR8 polypeptide, isolation may be accomplished using standard methods such as those described herein and in Marston *et al.*, 1990, *Meth. Enz.*, 182:264-75.

In some cases, an LGR8 polypeptide may not be biologically active upon isolation. Various methods for "refolding" or converting the polypeptide to its tertiary structure and generating disulfide linkages can be used to restore biological activity. Such methods include exposing the solubilized polypeptide to a pH usually above 7 and in the presence of a particular concentration of a chaotrope. The selection of chaotrope is very similar to the choices used for inclusion body solubilization, but usually the chaotrope is used at a lower concentration and is not necessarily the same as chaotropes used for the solubilization. In most cases the refolding/oxidation solution will also contain a reducing agent or the reducing agent plus its oxidized form in a specific ratio to generate a particular redox potential allowing for disulfide shuffling to occur in the formation of the protein's cysteine bridges. Some of the commonly used redox couples include cysteine/cystamine, glutathione (GSH)/dithiobis GSH, cupric chloride, dithiothreitol(DTT)/dithiane DTT, and 2-2-mercaptoethanol(bME)/dithio-b(ME). In many instances, a cosolvent may be used or may be needed to increase the efficiency of the refolding, and the more common reagents used for this purpose include glycerol, polyethylene glycol of various molecular weights, arginine and the like.

If inclusion bodies are not formed to a significant degree upon expression of an LGR8 polypeptide, then the polypeptide will be found primarily in the supernatant after centrifugation of the cell homogenate. The polypeptide may be further isolated from the supernatant using methods such as those described herein.

The purification of an LGR8 polypeptide from solution can be accomplished using a variety of techniques. If the polypeptide has been synthesized such that it contains a tag such as Hexahistidine (LGR8 polypeptide/hexaHis) or other small peptide such as FLAG (Eastman Kodak Co., New Haven, CT) or *myc* (Invitrogen) at either its carboxyl- or amino-terminus, it may be purified in a one-step process by passing the solution through an affinity column where the column matrix has a high affinity for the tag.

For example, polyhistidine binds with great affinity and specificity to nickel. Thus, an affinity column of nickel (such as the Qiagen® nickel columns) can be used

an mRNA and its encoded peptide. *See also,* Roberts, 1999, *Curr. Opin. Chem. Biol.* 3:268-73. Additionally, U.S. Patent No. 5,824,469 describes methods for obtaining oligonucleotides capable of carrying out a specific biological function. The procedure involves generating a heterogeneous pool of oligonucleotides, each having a 5' randomized sequence, a central preselected sequence, and a 3' randomized sequence. The resulting heterogeneous pool is introduced into a population of cells that do not exhibit the desired biological function. Subpopulations of the cells are then screened for those that exhibit a predetermined biological function. From that subpopulation, oligonucleotides capable of carrying out the desired biological function are isolated.

U.S. Patent Nos. 5,763,192; 5,814,476; 5,723,323; and 5,817,483 describe processes for producing peptides or polypeptides. This is done by producing stochastic genes or fragments thereof, and then introducing these genes into host cells which produce one or more proteins encoded by the stochastic genes. The host cells are then screened to identify those clones producing peptides or polypeptides having the desired activity.

Another method for producing peptides or polypeptides is described in PCT/US98/20094 (WO99/15650) filed by Athersys, Inc. Known as "Random Activation of Gene Expression for Gene Discovery" (RAGE-GD), the process involves the activation of endogenous gene expression or over-expression of a gene by *in situ* recombination methods. For example, expression of an endogenous gene is activated or increased by integrating a regulatory sequence into the target cell that is capable of activating expression of the gene by non-homologous or illegitimate recombination. The target DNA is first subjected to radiation, and a genetic promoter inserted. The promoter eventually locates a break at the front of a gene, initiating transcription of the gene. This results in expression of the desired peptide or polypeptide.

It will be appreciated that these methods can also be used to create comprehensive LGR8 polypeptide expression libraries, which can subsequently be used for high throughput phenotypic screening in a variety of assays, such as biochemical assays, cellular assays, and whole organism assays (*e.g.*, plant, mouse, etc.).

Synthesis

monoclonal antibodies include the hybridoma methods of Kohler *et al.*, 1975, *Nature* 256:495-97 and the human B-cell hybridoma method (Kozbor, 1984, *J. Immunol.* 133:3001; Brodeur *et al.*, *Monoclonal Antibody Production Techniques and Applications* 51-63 (Marcel Dekker, Inc., 1987). Also provided by the invention are hybridoma cell lines that produce monoclonal antibodies reactive with LGR8 polypeptides.

Monoclonal antibodies of the invention may be modified for use as therapeutics. One embodiment is a "chimeric" antibody in which a portion of the heavy (H) and/or light (L) chain is identical with or homologous to a corresponding sequence in antibodies derived from a particular species or belonging to a particular antibody class or subclass, while the remainder of the chain(s) is/are identical with or homologous to a corresponding sequence in antibodies derived from another species or belonging to another antibody class or subclass. Also included are fragments of such antibodies, so long as they exhibit the desired biological activity. *See* U.S. Patent No. 4,816,567; Morrison *et al.*, 1985, *Proc. Natl. Acad. Sci.* 81:6851-55.

In another embodiment, a monoclonal antibody of the invention is a "humanized" antibody. Methods for humanizing non-human antibodies are well known in the art. *See* U.S. Patent Nos. 5,585,089 and 5,693,762. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. Humanization can be performed, for example, using methods described in the art (Jones *et al.*, 1986, *Nature* 321:522-25; Riechmann *et al.*, 1998, *Nature* 332:323-27; Verhoeyen *et al.*, 1988, *Science* 239:1534-36), by substituting at least a portion of a rodent complementarity-determining region for the corresponding regions of a human antibody.

Also encompassed by the invention are human antibodies that bind LGR8 polypeptides. Using transgenic animals (*e.g.*, mice) that are capable of producing a repertoire of human antibodies in the absence of endogenous immunoglobulin production such antibodies are produced by immunization with an LGR8 polypeptide antigen (*i.e.*, having at least 6 contiguous amino acids), optionally conjugated to a carrier. *See, e.g.*, Jakobovits *et al.*, 1993, *Proc. Natl. Acad. Sci.* 90:2551-55; Jakobovits *et al.*, 1993, *Nature* 362:255-58; Bruggermann *et al.*, 1993, *Year in Immuno.* 7:33. In one method, such transgenic animals are produced by incapacitating the endogenous loci encoding the heavy and light immunoglobulin chains therein, and inserting loci encoding human heavy and light chain proteins into

the detectable moiety may be a radioisotope, such as ^{3}H, ^{14}C, ^{32}P, ^{35}S, ^{125}I, ^{99}Tc, ^{111}In, or ^{67}Ga; a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin; or an enzyme, such as alkaline phosphatase, β-galactosidase, or horseradish peroxidase (Bayer, *et al.*, 1990, *Meth. Enz.* 184:138-63).

Competitive binding assays rely on the ability of a labeled standard (*e.g.*, an LGR8 polypeptide, or an immunologically reactive portion thereof) to compete with the test sample analyte (an LGR8 polypeptide) for binding with a limited amount of anti-LGR8 antibody. The amount of an LGR8 polypeptide in the test sample is inversely proportional to the amount of standard that becomes bound to the antibodies. To facilitate determining the amount of standard that becomes bound, the antibodies typically are insolubilized before or after the competition, so that the standard and analyte that are bound to the antibodies may conveniently be separated from the standard and analyte that remain unbound.

Sandwich assays typically involve the use of two antibodies, each capable of binding to a different immunogenic portion, or epitope, of the protein to be detected and/or quantitated. In a sandwich assay, the test sample analyte is typically bound by a first antibody that is immobilized on a solid support, and thereafter a second antibody binds to the analyte, thus forming an insoluble three-part complex. *See, e.g.*, U.S. Patent No. 4,376,110. The second antibody may itself be labeled with a detectable moiety (direct sandwich assays) or may be measured using an anti-immunoglobulin antibody that is labeled with a detectable moiety (indirect sandwich assays). For example, one type of sandwich assay is an enzyme-linked immunosorbent assay (ELISA), in which case the detectable moiety is an enzyme.

The selective binding agents, including anti-LGR8 antibodies, are also useful for *in vivo* imaging. An antibody labeled with a detectable moiety may be administered to an animal, preferably into the bloodstream, and the presence and location of the labeled antibody in the host assayed. The antibody may be labeled with any moiety that is detectable in an animal, whether by nuclear magnetic resonance, radiology, or other detection means known in the art.

Selective binding agents of the invention, including antibodies, may be used as therapeutics. These therapeutic agents are generally agonists or antagonists, in that they either enhance or reduce, respectively, at least one of the biological activities of an LGR8 polypeptide. In one embodiment, antagonist antibodies of the invention are

of populations and generation of surrogate markers for clinical trials; and enhancing related LGR8 polypeptide small molecule drug discovery by aiding in the identification of selective compounds in high throughput screens.

Chemical Derivatives

Chemically modified derivatives of LGR8 polypeptides may be prepared by one skilled in the art, given the disclosures described herein. LGR8 polypeptide derivatives are modified in a manner that is different – either in the type or location of the molecules naturally attached to the polypeptide. Derivatives may include molecules formed by the deletion of one or more naturally-attached chemical groups. The polypeptide comprising the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or other LGR8 polypeptide, may be modified by the covalent attachment of one or more polymers. For example, the polymer selected is typically water-soluble so that the protein to which it is attached does not precipitate in an aqueous environment, such as a physiological environment. Included within the scope of suitable polymers is a mixture of polymers. Preferably, for therapeutic use of the end-product preparation, the polymer will be pharmaceutically acceptable.

The polymers each may be of any molecular weight and may be branched or unbranched. The polymers each typically have an average molecular weight of between about 2 kDa to about 100 kDa (the term "about" indicating that in preparations of a water-soluble polymer, some molecules will weigh more, some less, than the stated molecular weight). The average molecular weight of each polymer is preferably between about 5 kDa and about 50 kDa, more preferably between about 12 kDa and about 40 kDa and most preferably between about 20 kDa and about 35 kDa.

Suitable water-soluble polymers or mixtures thereof include, but are not limited to, N-linked or O-linked carbohydrates, sugars, phosphates, polyethylene glycol (PEG) (including the forms of PEG that have been used to derivatize proteins, including mono-(C_1-C_{10}), alkoxy-, or aryloxy-polyethylene glycol), monomethoxy-polyethylene glycol, dextran (such as low molecular weight dextran of, for example, about 6 kD), cellulose, or other carbohydrate based polymers, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene

(TNP) and the resulting conjugates precipitated with anti-DNP or anti-TNP-IgM to form decameric conjugates with a valency of 10.

Generally, conditions that may be alleviated or modulated by the administration of the present LGR8 polypeptide derivatives include those described herein for LGR8 polypeptides. However, the LGR8 polypeptide derivatives disclosed herein may have additional activities, enhanced or reduced biological activity, or other characteristics, such as increased or decreased half-life, as compared to the non-derivatized molecules.

Genetically Engineered Non-Human Animals

Additionally included within the scope of the present invention are non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which the genes encoding native LGR8 polypeptide have been disrupted (*i.e.*, "knocked out") such that the level of expression of LGR8 polypeptide is significantly decreased or completely abolished. Such animals may be prepared using techniques and methods such as those described in U.S. Patent No. 5,557,032.

The present invention further includes non-human animals such as mice, rats, or other rodents; rabbits, goats, sheep, or other farm animals, in which either the native form of an LGR8 gene for that animal or a heterologous LGR8 gene is over-expressed by the animal, thereby creating a "transgenic" animal. Such transgenic animals may be prepared using well known methods such as those described in U.S. Patent No 5,489,743 and PCT Pub. No. WO 94/28122.

The present invention further includes non-human animals in which the promoter for one or more of the LGR8 polypeptides of the present invention is either activated or inactivated (*e.g.*, by using homologous recombination methods) to alter the level of expression of one or more of the native LGR8 polypeptides.

These non-human animals may be used for drug candidate screening. In such screening, the impact of a drug candidate on the animal may be measured. For example, drug candidates may decrease or increase the expression of the LGR8 gene. In certain embodiments, the amount of LGR8 polypeptide that is produced may be measured after the exposure of the animal to the drug candidate. Additionally, in certain embodiments, one may detect the actual impact of the drug candidate on the animal. For example, over-expression of a particular gene may result in, or be associated with, a disease or pathological condition. In such cases, one may test a

sequences which direct or control the expression of LGR8 polypeptide, and which act as anti-sense regulators of expression.

Once a test molecule has been identified as interacting with an LGR8 polypeptide, the molecule may be further evaluated for its ability to increase or decrease LGR8 polypeptide activity. The measurement of the interaction of a test molecule with LGR8 polypeptide may be carried out in several formats, including cell-based binding assays, membrane binding assays, solution-phase assays, and immunoassays. In general, a test molecule is incubated with an LGR8 polypeptide for a specified period of time, and LGR8 polypeptide activity is determined by one or more assays for measuring biological activity.

The interaction of test molecules with LGR8 polypeptides may also be assayed directly using polyclonal or monoclonal antibodies in an immunoassay. Alternatively, modified forms of LGR8 polypeptides containing epitope tags as described herein may be used in solution and immunoassays.

In the event that LGR8 polypeptides display biological activity through an interaction with a binding partner (*e.g.*, a receptor or a ligand), a variety of *in vitro* assays may be used to measure the binding of an LGR8 polypeptide to the corresponding binding partner (such as a selective binding agent, receptor, or ligand). These assays may be used to screen test molecules for their ability to increase or decrease the rate and/or the extent of binding of an LGR8 polypeptide to its binding partner. In one assay, an LGR8 polypeptide is immobilized in the wells of a microtiter plate. Radiolabeled LGR8 polypeptide binding partner (for example, iodinated LGR8 polypeptide binding partner) and a test molecule can then be added either one at a time (in either order) or simultaneously to the wells. After incubation, the wells can be washed and counted for radioactivity, using a scintillation counter, to determine the extent to which the binding partner bound to the LGR8 polypeptide. Typically, a molecule will be tested over a range of concentrations, and a series of control wells lacking one or more elements of the test assays can be used for accuracy in the evaluation of the results. An alternative to this method involves reversing the "positions" of the proteins, *i.e.*, immobilizing LGR8 polypeptide binding partner to the microtiter plate wells, incubating with the test molecule and radiolabeled LGR8 polypeptide, and determining the extent of LGR8 polypeptide binding. *See*, *e.g.*, *Current Protocols in Molecular Biology*, chap. 18 (Ausubel *et al.*, eds., Green Publishers Inc. and Wiley and Sons 1995).

an LGR8 polypeptide and an LGR8 polypeptide binding partner. In these cases, the assays set forth herein can be readily modified by adding such additional test compound(s) either simultaneously with, or subsequent to, the first test compound. The remainder of the steps in the assay are as set forth herein.

In vitro assays such as those described herein may be used advantageously to screen large numbers of compounds for an effect on the formation of a complex between an LGR8 polypeptide and LGR8 polypeptide binding partner. The assays may be automated to screen compounds generated in phage display, synthetic peptide, and chemical synthesis libraries.

Compounds which increase or decrease the formation of a complex between an LGR8 polypeptide and an LGR8 polypeptide binding partner may also be screened in cell culture using cells and cell lines expressing either LGR8 polypeptide or LGR8 polypeptide binding partner. Cells and cell lines may be obtained from any mammal, but preferably will be from human or other primate, canine, or rodent sources. The binding of an LGR8 polypeptide to cells expressing LGR8 polypeptide binding partner at the surface is evaluated in the presence or absence of test molecules, and the extent of binding may be determined by, for example, flow cytometry using a biotinylated antibody to an LGR8 polypeptide binding partner. Cell culture assays can be used advantageously to further evaluate compounds that score positive in protein binding assays described herein.

Cell cultures can also be used to screen the impact of a drug candidate. For example, drug candidates may decrease or increase the expression of the LGR8 gene. In certain embodiments, the amount of LGR8 polypeptide or an LGR8 polypeptide fragment that is produced may be measured after exposure of the cell culture to the drug candidate. In certain embodiments, one may detect the actual impact of the drug candidate on the cell culture. For example, the over-expression of a particular gene may have a particular impact on the cell culture. In such cases, one may test a drug candidate's ability to increase or decrease the expression of the gene or its ability to prevent or inhibit a particular impact on the cell culture. In other examples, the production of a particular metabolic product such as a fragment of a polypeptide, may result in, or be associated with, a disease or pathological condition. In such cases, one may test a drug candidate's ability to decrease the production of such a metabolic product in a cell culture.

In accordance with certain embodiments of the invention, it may be useful to be able to determine the source of a certain cell type associated with an LGR8 polypeptide. For example, it may be useful to determine the origin of a disease or pathological condition as an aid in selecting an appropriate therapy. In certain embodiments, nucleic acids encoding an LGR8 polypeptide can be used as a probe to identify cells described herein by screening the nucleic acids of the cells with such a probe. In other embodiments, one may use anti-LGR8 polypeptide antibodies to test for the presence of LGR8 polypeptide in cells, and thus, determine if such cells are of the types described herein.

LGR8 Polypeptide Compositions and Administration

Therapeutic compositions are within the scope of the present invention. Such LGR8 polypeptide pharmaceutical compositions may comprise a therapeutically effective amount of an LGR8 polypeptide or an LGR8 nucleic acid molecule in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration. Pharmaceutical compositions may comprise a therapeutically effective amount of one or more LGR8 polypeptide selective binding agents in admixture with a pharmaceutically or physiologically acceptable formulation agent selected for suitability with the mode of administration.

Acceptable formulation materials preferably are nontoxic to recipients at the dosages and concentrations employed.

The pharmaceutical composition may contain formulation materials for modifying, maintaining, or preserving, for example, the pH, osmolarity, viscosity, clarity, color, isotonicity, odor, sterility, stability, rate of dissolution or release, adsorption, or penetration of the composition. Suitable formulation materials include, but are not limited to, amino acids (such as glycine, glutamine, asparagine, arginine, or lysine), antimicrobials, antioxidants (such as ascorbic acid, sodium sulfite, or sodium hydrogen-sulfite), buffers (such as borate, bicarbonate, Tris-HCl, citrates, phosphates, or other organic acids), bulking agents (such as mannitol or glycine), chelating agents (such as ethylenediamine tetraacetic acid (EDTA)), complexing agents (such as caffeine, polyvinylpyrrolidone, beta-cyclodextrin, or hydroxypropyl-beta-cyclodextrin), fillers, monosaccharides, disaccharides, and other carbohydrates (such as glucose, mannose, or dextrins), proteins (such as serum albumin, gelatin, or immunoglobulins), coloring, flavoring and diluting agents, emulsifying agents,

pharmaceutically acceptable compositions is within the skill of the art.

The formulation components are present in concentrations that are acceptable to the site of administration. For example, buffers are used to maintain the composition at physiological pH or at a slightly lower pH, typically within a pH range of from about 5 to about 8.

When parenteral administration is contemplated, the therapeutic compositions for use in this invention may be in the form of a pyrogen-free, parenterally acceptable, aqueous solution comprising the desired LGR8 molecule in a pharmaceutically acceptable vehicle. A particularly suitable vehicle for parenteral injection is sterile distilled water in which an LGR8 molecule is formulated as a sterile, isotonic solution, properly preserved. Yet another preparation can involve the formulation of the desired molecule with an agent, such as injectable microspheres, bio-erodible particles, polymeric compounds (such as polylactic acid or polyglycolic acid), beads, or liposomes, that provides for the controlled or sustained release of the product which may then be delivered via a depot injection. Hyaluronic acid may also be used, and this may have the effect of promoting sustained duration in the circulation. Other suitable means for the introduction of the desired molecule include implantable drug delivery devices.

In one embodiment, a pharmaceutical composition may be formulated for inhalation. For example, LGR8 polypeptide may be formulated as a dry powder for inhalation. LGR8 polypeptide or nucleic acid molecule inhalation solutions may also be formulated with a propellant for aerosol delivery. In yet another embodiment, solutions may be nebulized. Pulmonary administration is further described in PCT Pub. No. WO 94/20069, which describes the pulmonary delivery of chemically modified proteins.

It is also contemplated that certain formulations may be administered orally. In one embodiment of the present invention, LGR8 polypeptides that are administered in this fashion can be formulated with or without those carriers customarily used in the compounding of solid dosage forms such as tablets and capsules. For example, a capsule may be designed to release the active portion of the formulation at the point in the gastrointestinal tract when bioavailability is maximized and pre-systemic degradation is minimized. Additional agents can be included to facilitate absorption of the LGR8 polypeptide. Diluents, flavorings, low melting point waxes, vegetable

lyophilized form or in a solution. In addition, parenteral compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Once the pharmaceutical composition has been formulated, it may be stored in sterile vials as a solution, suspension, gel, emulsion, solid, or as a dehydrated or lyophilized powder. Such formulations may be stored either in a ready-to-use form or in a form (*e.g.*, lyophilized) requiring reconstitution prior to administration.

In a specific embodiment, the present invention is directed to kits for producing a single-dose administration unit. The kits may each contain both a first container having a dried protein and a second container having an aqueous formulation. Also included within the scope of this invention are kits containing single and multi-chambered pre-filled syringes (*e.g.*, liquid syringes and lyosyringes).

The effective amount of an LGR8 pharmaceutical composition to be employed therapeutically will depend, for example, upon the therapeutic context and objectives. One skilled in the art will appreciate that the appropriate dosage levels for treatment will thus vary depending, in part, upon the molecule delivered, the indication for which the LGR8 molecule is being used, the route of administration, and the size (body weight, body surface, or organ size) and condition (the age and general health) of the patient. Accordingly, the clinician may titer the dosage and modify the route of administration to obtain the optimal therapeutic effect. A typical dosage may range from about 0.1 μg/kg to up to about 100 mg/kg or more, depending on the factors mentioned above. In other embodiments, the dosage may range from 0.1 μg/kg up to about 100 mg/kg; or 1 μg/kg up to about 100 mg/kg; or 5 μg/kg up to about 100 mg/kg.

The frequency of dosing will depend upon the pharmacokinetic parameters of the LGR8 molecule in the formulation being used. Typically, a clinician will administer the composition until a dosage is reached that achieves the desired effect. The composition may therefore be administered as a single dose, as two or more doses (which may or may not contain the same amount of the desired molecule) over time, or as a continuous infusion via an implantation device or catheter. Further refinement of the appropriate dosage is routinely made by those of ordinary skill in the art and is within the ambit of tasks routinely performed by them. Appropriate dosages may be ascertained through use of appropriate dose-response data.

polypeptides.

Homologous recombination is a technique originally developed for targeting genes to induce or correct mutations in transcriptionally active genes. Kucherlapati, 1989, *Prog. in Nucl. Acid Res. & Mol. Biol.* 36:301. The basic technique was developed as a method for introducing specific mutations into specific regions of the mammalian genome (Thomas *et al.*, 1986, *Cell* 44:419-28; Thomas and Capecchi, 1987, *Cell* 51:503-12; Doetschman *et al.*, 1988, *Proc. Natl. Acad. Sci. U.S.A.* 85:8583-87) or to correct specific mutations within defective genes (Doetschman *et al.*, 1987, *Nature* 330:576-78). Exemplary homologous recombination techniques are described in U.S. Patent No. 5,272,071; European Patent Nos. 9193051 and 505500; PCT/US90/07642, and PCT Pub No. WO 91/09955).

Through homologous recombination, the DNA sequence to be inserted into the genome can be directed to a specific region of the gene of interest by attaching it to targeting DNA. The targeting DNA is a nucleotide sequence that is complementary (homologous) to a region of the genomic DNA. Small pieces of targeting DNA that are complementary to a specific region of the genome are put in contact with the parental strand during the DNA replication process. It is a general property of DNA that has been inserted into a cell to hybridize, and therefore, recombine with other pieces of endogenous DNA through shared homologous regions. If this complementary strand is attached to an oligonucleotide that contains a mutation or a different sequence or an additional nucleotide, it too is incorporated into the newly synthesized strand as a result of the recombination. As a result of the proofreading function, it is possible for the new sequence of DNA to serve as the template. Thus, the transferred DNA is incorporated into the genome.

Attached to these pieces of targeting DNA are regions of DNA that may interact with or control the expression of an LGR8 polypeptide, *e.g.*, flanking sequences. For example, a promoter/enhancer element, a suppressor, or an exogenous transcription modulatory element is inserted in the genome of the intended host cell in proximity and orientation sufficient to influence the transcription of DNA encoding the desired LGR8 polypeptide. The control element controls a portion of the DNA present in the host cell genome. Thus, the expression of the desired LGR8 polypeptide may be achieved not by transfection of DNA that encodes the LGR8 gene itself, but rather by the use of targeting DNA (containing regions of homology with the endogenous gene of interest) coupled with DNA regulatory segments that provide

properly positioned in this plasmid, would integrate in such a manner as to create a new or modified transcriptional unit resulting in *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene.

A further method to use the cell line in which the site specific recombination sequence had been placed just upstream of the cell's endogenous genomic LGR8 polypeptide coding region is to use homologous recombination to introduce a second recombination site elsewhere in the cell line's genome. The appropriate recombinase enzyme is then introduced into the two-recombination-site cell line, causing a recombination event (deletion, inversion, and translocation) (Sauer, 1994, *Curr. Opin. Biotechnol.*, 5:521-27; Sauer, 1993, *Methods Enzymol.*, 225:890-900) that would create a new or modified transcriptional unit resulting in *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene.

An additional approach for increasing, or causing, the expression of LGR8 polypeptide from a cell's endogenous LGR8 gene involves increasing, or causing, the expression of a gene or genes (*e.g.*, transcription factors) and/or decreasing the expression of a gene or genes (*e.g.*, transcriptional repressors) in a manner which results in *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene. This method includes the introduction of a non-naturally occurring polypeptide (*e.g.*, a polypeptide comprising a site specific DNA binding domain fused to a transcriptional factor domain) into the cell such that *de novo* or increased LGR8 polypeptide production from the cell's endogenous LGR8 gene results.

The present invention further relates to DNA constructs useful in the method of altering expression of a target gene. In certain embodiments, the exemplary DNA constructs comprise: (a) one or more targeting sequences, (b) a regulatory sequence, (c) an exon, and (d) an unpaired splice-donor site. The targeting sequence in the DNA construct directs the integration of elements (a) - (d) into a target gene in a cell such that the elements (b) - (d) are operatively linked to sequences of the endogenous target gene. In another embodiment, the DNA constructs comprise: (a) one or more targeting sequences, (b) a regulatory sequence, (c) an exon, (d) a splice-donor site, (e) an intron, and (f) a splice-acceptor site, wherein the targeting sequence directs the integration of elements (a) - (f) such that the elements of (b) - (f) are operatively linked to the endogenous gene. The targeting sequence is homologous to the preselected site in the cellular chromosomal DNA with which homologous

Alternatively, the patient's own cells, transformed to produce LGR8 polypeptides *ex vivo*, may be implanted directly into the patient without such encapsulation.

Techniques for the encapsulation of living cells are known in the art, and the preparation of the encapsulated cells and their implantation in patients may be routinely accomplished. For example, Baetge *et al.* (PCT Pub. No. WO 95/05452 and PCT/US94/09299) describe membrane capsules containing genetically engineered cells for the effective delivery of biologically active molecules. The capsules are biocompatible and are easily retrievable. The capsules encapsulate cells transfected with recombinant DNA molecules comprising DNA sequences coding for biologically active molecules operatively linked to promoters that are not subject to down-regulation *in vivo* upon implantation into a mammalian host. The devices provide for the delivery of the molecules from living cells to specific sites within a recipient. In addition, *see* U.S. Patent Nos. 4,892,538; 5,011,472; and 5,106,627. A system for encapsulating living cells is described in PCT Pub. No. WO 91/10425 (Aebischer *et al.*). *See also*, PCT Pub. No. WO 91/10470 (Aebischer *et al.*); Winn *et al.*, 1991, *Exper. Neurol.* 113:322-29; Aebischer *et al.*, 1991, *Exper. Neurol.* 111:269-75; and Tresco *et al.*, 1992, *ASAIO* 38:17-23.

In vivo and *in vitro* gene therapy delivery of LGR8 polypeptides is also envisioned. One example of a gene therapy technique is to use the LGR8 gene (either genomic DNA, cDNA, and/or synthetic DNA) encoding an LGR8 polypeptide that may be operably linked to a constitutive or inducible promoter to form a "gene therapy DNA construct." The promoter may be homologous or heterologous to the endogenous LGR8 gene, provided that it is active in the cell or tissue type into which the construct will be inserted. Other components of the gene therapy DNA construct may optionally include DNA molecules designed for site-specific integration (*e.g.*, endogenous sequences useful for homologous recombination), tissue-specific promoters, enhancers or silencers, DNA molecules capable of providing a selective advantage over the parent cell, DNA molecules useful as labels to identify transformed cells, negative selection systems, cell specific binding agents (as, for example, for cell targeting), cell-specific internalization factors, transcription factors enhancing expression from a vector, and factors enabling vector production.

A gene therapy DNA construct can then be introduced into cells (either *ex vivo* or *in vivo*) using viral or non-viral vectors. One means for introducing the gene therapy DNA construct is by means of viral vectors as described herein. Certain

ecdysone-responsive gene). The ecdysone receptor includes a transactivation domain, DNA-binding domain, and ligand-binding domain to initiate transcription. The ecdysone system is further described in U.S. Patent No. 5,514,578 and PCT Pub. Nos. WO 97/38117, WO 96/37609, and WO 93/03162.

Another control means uses a positive tetracycline-controllable transactivator. This system involves a mutated tet repressor protein DNA-binding domain (mutated tet R-4 amino acid changes which resulted in a reverse tetracycline-regulated transactivator protein, *i.e.*, it binds to a tet operator in the presence of tetracycline) linked to a polypeptide which activates transcription. Such systems are described in U.S. Patent Nos. 5,464,758, 5,650,298, and 5,654,168.

Additional expression control systems and nucleic acid constructs are described in U.S. Patent Nos. 5,741,679 and 5,834,186, to Innovir Laboratories Inc.

In vivo gene therapy may be accomplished by introducing the gene encoding LGR8 polypeptide into cells via local injection of an LGR8 nucleic acid molecule or by other appropriate viral or non-viral delivery vectors. Hefti 1994, *Neurobiology* 25:1418-35. For example, a nucleic acid molecule encoding an LGR8 polypeptide may be contained in an adeno-associated virus (AAV) vector for delivery to the targeted cells (*see, e.g.*, Johnson, PCT Pub. No. WO 95/34670; PCT App. No. PCT/US95/07178). The recombinant AAV genome typically contains AAV inverted terminal repeats flanking a DNA sequence encoding an LGR8 polypeptide operably linked to functional promoter and polyadenylation sequences.

Alternative suitable viral vectors include, but are not limited to, retrovirus, adenovirus, herpes simplex virus, lentivirus, hepatitis virus, parvovirus, papovavirus, poxvirus, alphavirus, coronavirus, rhabdovirus, paramyxovirus, and papilloma virus vectors. U.S. Patent No. 5,672,344 describes an *in vivo* viral-mediated gene transfer system involving a recombinant neurotrophic HSV-1 vector. U.S. Patent No. 5,399,346 provides examples of a process for providing a patient with a therapeutic protein by the delivery of human cells that have been treated *in vitro* to insert a DNA segment encoding a therapeutic protein. Additional methods and materials for the practice of gene therapy techniques are described in U.S. Patent Nos. 5,631,236 (involving adenoviral vectors), 5,672,510 (involving retroviral vectors), 5,635,399 (involving retroviral vectors expressing cytokines).

Nonviral delivery methods include, but are not limited to, liposome-mediated transfer, naked DNA delivery (direct injection), receptor-mediated transfer (ligand-

Gene therapy also can be used to decrease LGR8 polypeptide expression by modifying the nucleotide sequence of the endogenous promoter. Such modification is typically accomplished via homologous recombination methods. For example, a DNA molecule containing all or a portion of the promoter of the LGR8 gene selected for inactivation can be engineered to remove and/or replace pieces of the promoter that regulate transcription. For example, the TATA box and/or the binding site of a transcriptional activator of the promoter may be deleted using standard molecular biology techniques; such deletion can inhibit promoter activity thereby repressing the transcription of the corresponding LGR8 gene. The deletion of the TATA box or the transcription activator binding site in the promoter may be accomplished by generating a DNA construct comprising all or the relevant portion of the LGR8 polypeptide promoter (from the same or a related species as the LGR8 gene to be regulated) in which one or more of the TATA box and/or transcriptional activator binding site nucleotides are mutated via substitution, deletion and/or insertion of one or more nucleotides. As a result, the TATA box and/or activator binding site has decreased activity or is rendered completely inactive. This construct, which also will typically contain at least about 500 bases of DNA that correspond to the native (endogenous) 5' and 3' DNA sequences adjacent to the promoter segment that has been modified, may be introduced into the appropriate cells (either *ex vivo* or *in vivo*) either directly or via a viral vector as described herein. Typically, the integration of the construct into the genomic DNA of the cells will be via homologous recombination, where the 5' and 3' DNA sequences in the promoter construct can serve to help integrate the modified promoter region via hybridization to the endogenous chromosomal DNA.

Therapeutic Uses

LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof can be used to treat, diagnose, ameliorate, or prevent a number of diseases, disorders, or conditions, including those recited herein.

LGR8 polypeptide agonists and antagonists include those molecules which regulate LGR8 polypeptide activity and either increase or decrease at least one activity of the mature form of the LGR8 polypeptide. Agonists or antagonists may be co-factors, such as a protein, peptide, carbohydrate, lipid, or small molecular weight molecule, which interact with LGR8 polypeptide and thereby regulate its activity.

Since LGR8 polypeptide expression has been detected in the uterus, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting skeletal muscle. Examples of such diseases and conditions include, but are not limited to, miscarriage, endometriosis, uterine cancer, and female infertility. Other diseases and conditions associated with uterine development and function are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the adrenal gland, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the adrenal gland. Examples of such diseases and conditions include, but are not limited to, Cushing's disease and Addison's disease. Other diseases and conditions associated with the development and function of the adrenal gland are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the testes, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the testes. Examples of such diseases and conditions include, but are not limited to, male infertility and testicular carcinoma. Other diseases and conditions associated with the development and function of the testes are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the bone marrow, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the bone marrow. Examples of such diseases and conditions include, but are not limited to, leukemia. Other diseases and conditions associated with the development and function of the bone marrow are encompassed within the scope of this invention.

Since LGR8 polypeptide expression has been detected in the fetal kidney, LGR8 nucleic acid molecules, polypeptides, and agonists and antagonists thereof may be useful in diagnosing or treating diseases and conditions affecting the kidney. Examples of such diseases and conditions include, but are not limited to, anemia, hypertension, and low blood pressure. Other diseases and conditions associated with the development and function of the kidney are encompassed within the scope of this invention.

Nucleic acid molecules of the invention (including those that do not themselves encode biologically active polypeptides) may be used to map the locations of the LGR8 gene and related genes on chromosomes. Mapping may be done by techniques known in the art, such as PCR amplification and *in situ* hybridization.

LGR8 nucleic acid molecules (including those that do not themselves encode biologically active polypeptides), may be useful as hybridization probes in diagnostic assays to test, either qualitatively or quantitatively, for the presence of an LGR8 nucleic acid molecule in mammalian tissue or bodily fluid samples.

Other methods may also be employed where it is desirable to inhibit the activity of one or more LGR8 polypeptides. Such inhibition may be effected by nucleic acid molecules that are complementary to and hybridize to expression control sequences (triple helix formation) or to LGR8 mRNA. For example, antisense DNA or RNA molecules, which have a sequence that is complementary to at least a portion of an LGR8 gene can be introduced into the cell. Anti-sense probes may be designed by available techniques using the sequence of the LGR8 gene disclosed herein. Typically, each such antisense molecule will be complementary to the start site (5' end) of each selected LGR8 gene. When the antisense molecule then hybridizes to the corresponding LGR8 mRNA, translation of this mRNA is prevented or reduced. Anti-sense inhibitors provide information relating to the decrease or absence of an LGR8 polypeptide in a cell or organism.

Alternatively, gene therapy may be employed to create a dominant-negative inhibitor of one or more LGR8 polypeptides. In this situation, the DNA encoding a mutant polypeptide of each selected LGR8 polypeptide can be prepared and introduced into the cells of a patient using either viral or non-viral methods as described herein. Each such mutant is typically designed to compete with endogenous polypeptide in its biological role.

In addition, an LGR8 polypeptide, whether biologically active or not, may be used as an immunogen, that is, the polypeptide contains at least one epitope to which antibodies may be raised. Selective binding agents that bind to an LGR8 polypeptide (as described herein) may be used for *in vivo* and *in vitro* diagnostic purposes, including, but not limited to, use in labeled form to detect the presence of LGR8 polypeptide in a body fluid or cell sample. The antibodies may also be used to prevent, treat, or diagnose a number of diseases and disorders, including those recited herein. The antibodies may bind to an LGR8 polypeptide so as to diminish or block

encoding the mature form (*i.e.*, lacking the signal peptide) of human LGR8. These sequences were then used to design gene specific oligonucleotides for the identification of cDNA sources and the generation of cDNA clones, using various PCR strategies. Several highly homologous, but not identical, LGR8 sequences were thus isolated. An analysis of these sequences led to the identification of four nucleotide sequences encoding the mature forms (*i.e.*, the forms lacking an LGR8 start codon and the nucleotide sequence encoding the signal peptide) of LGR8-A, LGR8-B, LGR8-C, and LGR8-D.

The nucleotide sequence encoding the mature form of human LGR8-A was obtained in amplification reactions using 5 µl of a human adrenal Marathon Ready cDNA template (Clontech Laboratories; Palo Alto, CA), 1.0 µm each of the amplimers 5'-T-G-C-C-A-A-A-A-A-G-G-A-T-A-T-T-T-T-C-C-C-T-G-T-G-G-G-A-A-T-C-T-T-A-3' (SEQ ID NO: 27) and 5'-C-T-A-G-G-A-AA-C-T-G-G-T-T-T-C-A-T-T-A-T-A-C-T-G-T-C-T-C-C-A-A-G-T-G-T-T-A-T-T-T-T-G-T-T-C-A-3' (SEQ ID NO: 28), 200 µm of dNTPs, 2.5 U of *Pfu*Turbo DNA polymerase (Stratagene; La Jolla, CA), and 5 µL 10X *Pfu*Turbo DNA polymerase reaction buffer in a final volume of 50 µL. Reactions were performed at 94°C for 1 minute for one cycle; 94°C for 10 seconds, 60°C for 20 seconds, and 72°C for 5 minutes for 45 cycles; and 72°C for 7 minutes for one cycle.

The amplification mixture was separated on an agarose gel, the PCR products were isolated from the gel, and the products were then blunt-end cloned into pPCR-Script Amp SK(+) (Stratagene). A number of clones were sequenced, each containing highly homologous (but not identical) LGR8 nucleotide sequences. These sequences were used to compile the four nucleotide sequences encoding the mature forms (*i.e.*, the form lacking a start codon and a signal peptide) of LGR8-A, LGR8-B, LGR8-C, and LGR8-D.

To isolate cDNA sequences corresponding to the 5' end of the cDNA sequence for the immature form of human LGR8, 5' RACE was performed using 5 µL of a human adrenal Marathon Ready cDNA template, 1.0 µm each of the primers 5'-C-C-A-T-C-C-T-A-A-T-A-C-G-A-C-T-C-A-C-T-A-T-A-G-G-G-C-3' (SEQ ID NO: 29) and 5'-A-T-T-G-T-C-A-T-C-T-A-G-A-A-T-T-A-G-C-C-A-A-G-T-T-A-G-C-T-G-A-T-3' (SEQ ID NO: 30), 200 µm of dNTPs, 1 µL of 50X Advantage2 Polymerase Mix (Clontech Laboratories), and 5 µL 10X Advantage2 PCR buffer in a final volume of 50 µL. Reactions were performed at 94°C for 1 minute for one cycle;

cytoplasmic C-terminal region. Sequence analysis of the full-length LGR8-A coding sequence indicated that the cDNA comprises a 2262 bp open reading frame encoding a protein of 754 amino acids (Figures 1A-1D). The mature form of LGR8-A is 718 amino acids in length. LGR8-A is most closely related to glycoprotein hormone receptor LGR7 (Figures 10A-10B).

The LGR8-B coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-B coding sequence lacks a portion of the sequence encoding the N-terminal extracellular domain. Sequence analysis of the full-length LGR8-B coding sequence indicated that the cDNA comprises a 2190 bp open reading frame encoding a protein of 730 amino acids (Figures 3A-3D). The mature form of LGR8-B is 694 amino acids in length.

The LGR8-C coding sequence is identical to the LGR8-A coding sequence with the exception that the LGR8-C coding sequence lacks a portion of sequence encoding the N-terminal extracellular domain. Sequence analysis of the full-length LGR8-C coding sequence indicated that the cDNA comprises a 2046 bp open reading frame encoding a protein of 682 amino acids (Figures 5A-5D). The mature form of LGR8-C is 646 amino acids in length.

The LGR8-D coding sequence consists of sequence encoding approximately 90% of the N-terminal extracellular domain of the LGR8-B coding sequence, but lacks sequence encoding the transmembrane domains and the cytoplasmic C-terminal region. Sequence analysis of the full-length LGR8-D coding sequence indicated that the cDNA comprises a 1098 bp open reading frame encoding a protein of 366 amino acids (Figures 7A-7B). The mature form of LGR8-D is 330 amino acids in length.

To identify cDNA sequences encoding murine LGR8-A, homology-based BLAST searches of a human genomic database were performed using the amino acid sequence of human LGR8-A. A number of sequences sharing a high degree of homology were found within a 213 kb mouse public genomic sequence (Accession No. AC077689). No exons, genes, or homologies to known genes were identified in the AC077689 sequence record. The sequences thus identified were currated by hand and electronically compiled to yield the complete nucleotide sequence encoding murine LGR8-A (Figures 8A-8D). A sequence comparison of the mature human and murine LGR8-A sequences indicates that the sequences share an 86.6% similarity and an 83.1% identity (Figures 11A-11B). A sequence comparison of the N-terminal extracellular domains (absent the signal peptide) of human and murine LGR8-A

an agarose gel, and PCR products of the expected size (319 bp) were identified in adult skeletal muscle, testis, and adrenal gland.

Intron-spanning PCR was next performed on proprietary oligo-dT primed and random primed human cDNA libraries for the following tissues: fetal stomach (oligo-dT primed), fetal stomach (random primed), pons/medulla (oligo-dT primed), breast tumor T1485 (oligo-dT primed), breast tumor T1485 (random primed), ovary tumor T22 (oligo-dT primed), ovary tumor T22 (random primed), fetal thymus (oligo-dT primed), fetal thymus (random primed), fetal mesentery (oligo-dT primed), fetal mesentery (random primed), placenta (oligo-dT primed), placenta (random primed), Multiple cell lines [A204, A673, Hs729T, HISM and RD (oligo-dT primed)], Multiple cell lines [A204, A673, Hs729T, HISM and RD (random primed)], fetal pancreas (oligo-dT primed), fetal pancreas (random primed), lymphoma cell lines (oligo-dT primed), lymphoma cell lines (random primed), ovary tumor T23 (oligo-dT primed), ovary tumor T23 (random primed), colon tumor T25 (oligo-dT primed), colon tumor T25 (random primed), adult T-cells (oligo-dT primed), normalized fetal tissue (random primed), fetal heart (oligo-dT primed), fetal heart (random primed), fetal bladder (oligo-dT primed), fetal bladder (random primed), fetal kidney (oligo-dT primed), fetal kidney (random primed), lung tumor T27 (oligo-dT primed), lung tumor T27 (random primed), fetal liver (oligo-dT primed), cytoplasmic breast carcinoma cell lines (oligo-dT primed), cytoplasmic breast carcinoma cell lines (random primed), fetal spleen (oligo-dT primed), fetal spleen (random primed), uterus (oligo-dT primed), uterus (random primed), adrenal (oligo-dT primed), adrenal (random primed), forebrain (oligo-dT primed), forebrain (random primed), testis (oligo-dT primed), testis (random primed), colon tumor T24 (oligo-dT primed), colon tumor T24 (random primed), fetal heart (oligo-dT primed), fetal scalp (oligo-dT primed), fetal scalp (random primed), fetal lung (oligo-dT primed), fetal lung (random primed), trachea (oligo-dT primed), trachea (random primed), cerebellum (oligo-dT primed), midbrain LNV block 10 (oligo-dT primed), midbrain LNV block 10 (random primed), prostate tumor T1940 (random primed), fetal ovary (oligo-dT primed), fetal calveria (oligo-dT primed), fetal calveria (random primed), fetal gall bladder (oligo-dT primed), fetal gall bladder (random primed), spinal column (oligo-dT primed), spinal column (random primed), thalamus (oligo-dT primed), prostate tumor T1175 (oligo-dT primed), prostate tumor T1175 (random primed), breast tumor T1543 (oligo-dT primed), breast tumor T1543 (random primed), fetal skin (oligo-dT

performed as described above using 300 nm each of the primers 5'-G-T-C-G-A-C-G-G-C-G-A-G-C-C-C-3' (SEQ ID NO: 40) and 5'-T-C-T-T-T-G-G-G-A-C-C-T-T-G-T-C-T-G-C-A-A-3' (SEQ ID NO: 41) and 200 nm of the flurogenic probe 5'-(6-FAM)-T-G-G-G-C-C-G-C-G-T-C-T-C-C-T-T-T-G-A-G-C-T-(TAMRA)-3' (Primer Express, PE BioSystems, Foster City, CA; SEQ ID NO: 42; wherein "6-FAM" is the 5' reporter dye 6-carboxy-fluorescein and "TAMRA" is the 3' quencher 6-carboxytetramethylrhodamine). The highest levels of LGR8 mRNA expression were detected in skeletal muscle and uterus. Lower levels were found in adrenal and testis, with lower levels still in thalamus and bone marrow.

The expression of LGR8 mRNA is localized by *in situ* hybridization. A panel of normal embryonic and adult mouse tissues is fixed in 4% paraformaldehyde, embedded in paraffin, and sectioned at 5 μm. Sectioned tissues are permeabilized in 0.2 M HCl, digested with Proteinase K, and acetylated with triethanolamine and acetic anhydride. Sections are prehybridized for 1 hour at 60°C in hybridization solution (300 mM NaCl, 20 mM Tris-HCl, pH 8.0, 5 mM EDTA, 1X Denhardt's solution, 0.2% SDS, 10 mM DTT, 0.25 mg/ml tRNA, 25 μg/ml polyA, 25 μg/ml polyC and 50% formamide) and then hybridized overnight at 60°C in the same solution containing 10% dextran and 2 x 10^4 cpm/μl of a ^{33}P-labeled antisense riboprobe complementary to the human LGR8 gene. The riboprobe is obtained by *in vitro* transcription of a clone containing human LGR8 cDNA sequences using standard techniques.

Following hybridization, sections are rinsed in hybridization solution, treated with RNaseA to digest unhybridized probe, and then washed in 0.1X SSC at 55°C for 30 minutes. Sections are then immersed in NTB-2 emulsion (Kodak, Rochester, NY), exposed for 3 weeks at 4°C, developed, and counterstained with hematoxylin and eosin. Tissue morphology and hybridization signal are simultaneously analyzed by darkfield and standard illumination for brain (one sagittal and two coronal sections), gastrointestinal tract (esophagus, stomach, duodenum, jejunum, ileum, proximal colon, and distal colon), pituitary, liver, lung, heart, spleen, thymus, lymph nodes, kidney, adrenal, bladder, pancreas, salivary gland, male and female reproductive organs (ovary, oviduct, and uterus in the female; and testis, epididymus, prostate, seminal vesicle, and vas deferens in the male), BAT and WAT (subcutaneous, perirenal), bone (femur), skin, breast, and skeletal muscle.

B. Expression of LGR8 Polypeptide in Mammalian Cells

PCR is used to amplify template DNA sequences encoding an LGR8 polypeptide using primers corresponding to the 5' and 3' ends of the sequence. The amplified DNA products may be modified to contain restriction enzyme sites to allow for insertion into expression vectors. PCR products are gel purified and inserted into expression vectors using standard recombinant DNA methodology. An exemplary expression vector, pCEP4 (Invitrogen, Carlsbad, CA), that contains an Epstein-Barr virus origin of replication, may be used for the expression of LGR8 polypeptides in 293-EBNA-1 cells. Amplified and gel purified PCR products are ligated into pCEP4 vector and introduced into 293-EBNA cells by lipofection. The transfected cells are selected in 100 μg/mL hygromycin and the resulting drug-resistant cultures are grown to confluence. The cells are then cultured in serum-free media for 72 hours. The conditioned media is removed and LGR8 polypeptide expression is analyzed by SDS-PAGE.

LGR8 polypeptide expression may be detected by silver staining. Alternatively, LGR8 polypeptide is produced as a fusion protein with an epitope tag, such as an IgG constant domain or a FLAG epitope, which may be detected by Western blot analysis using antibodies to the peptide tag.

LGR8 polypeptides may be excised from an SDS-polyacrylamide gel, or LGR8 fusion proteins are purified by affinity chromatography to the epitope tag, and subjected to N-terminal amino acid sequence analysis as described herein.

C. Purification of LGR8 Polypeptide from Mammalian Cells

LGR8 polypeptide expression constructs are introduced into 293 EBNA or CHO cells using either a lipofection or calcium phosphate protocol.

To conduct functional studies on the LGR8 polypeptides that are produced, large quantities of conditioned media are generated from a pool of hygromycin selected 293 EBNA clones. The cells are cultured in 500 cm Nunc Triple Flasks to 80% confluence before switching to serum-free media a week prior to harvesting the media. Conditioned media is harvested and frozen at
-20°C until the protein is to be purified.

Conditioned media is purified by affinity chromatography as described below. The media is thawed and then passed through a 0.2 μm filter. A Protein G column is equilibrated with PBS at pH 7.0, and then loaded with the filtered media. The column

changes that are informative as to the function of LGR8 polypeptide. Similarly, a construct containing the full-length LGR8 polypeptide under the control of the beta actin promoter is prepared. The delivery of this construct is expected to result in ubiquitous expression.

To generate these constructs, PCR is used to amplify template DNA sequences encoding an LGR8 polypeptide using primers that correspond to the 5' and 3' ends of the desired sequence and which incorporate restriction enzyme sites to permit insertion of the amplified product into an expression vector. Following amplification, PCR products are gel purified, digested with the appropriate restriction enzymes, and ligated into an expression vector using standard recombinant DNA techniques. For example, amplified LGR8 polypeptide sequences can be cloned into an expression vector under the control of the human β-actin promoter as described by Graham *et al.*, 1997, *Nature Genetics*, 17:272-74 and Ray *et al.*, 1991, *Genes Dev.* 5:2265-73.

Following ligation, reaction mixtures are used to transform an *E. coli* host strain by electroporation and transformants are selected for drug resistance. Plasmid DNA from selected colonies is isolated and subjected to DNA sequencing to confirm the presence of an appropriate insert and absence of mutation. The LGR8 polypeptide expression vector is purified through two rounds of CsCl density gradient centrifugation, cleaved with a suitable restriction enzyme, and the linearized fragment containing the LGR8 polypeptide transgene is purified by gel electrophoresis. The purified fragment is resuspended in 5 mM Tris, pH 7.4, and 0.2 mM EDTA at a concentration of 2 mg/mL.

Single-cell embryos from BDF1 x BDF1 bred mice are injected as described (PCT Pub. No. WO 97/23614). Embryos are cultured overnight in a CO_2 incubator and 15-20 two-cell embryos are transferred to the oviducts of a pseudopregnant CD1 female mice. Offspring obtained from the implantation of microinjected embryos are screened by PCR amplification of the integrated transgene in genomic DNA samples as follows. Ear pieces are digested in 20 mL ear buffer (20 mM Tris, pH 8.0, 10 mM EDTA, 0.5% SDS, and 500 mg/mL proteinase K) at 55°C overnight. The sample is then diluted with 200 mL of TE, and 2 mL of the ear sample is used in a PCR reaction using appropriate primers.

At 8 weeks of age, transgenic founder animals and control animals are sacrificed for necropsy and pathological analysis. Portions of spleen are removed and total cellular RNA isolated from the spleens using the Total RNA Extraction Kit

chromagen (BioTek, Santa Barbara, CA). Sections are counterstained with hematoxylin.

After necropsy, MLN and sections of spleen and thymus from transgenic animals and control littermates are removed. Single cell suspensions are prepared by gently grinding the tissues with the flat end of a syringe against the bottom of a 100 mm nylon cell strainer (Becton Dickinson, Franklin Lakes, NJ). Cells are washed twice, counted, and approximately 1 x 10^6 cells from each tissue are then incubated for 10 minutes with 0.5 μg CD16/32(FcγIII/II) Fc block in a 20 μL volume. Samples are then stained for 30 minutes at 2-8°C in a 100 μL volume of PBS (lacking Ca^+ and Mg^+), 0.1% bovine serum albumin, and 0.01% sodium azide with 0.5 μg antibody of FITC or PE-conjugated monoclonal antibodies against CD90.2 (Thy-1.2), CD45R (B220), CD11b(Mac-1), Gr-1, CD4, or CD8 (PharMingen, San Diego, CA). Following antibody binding, the cells are washed and then analyzed by flow cytometry on a FACScan (Becton Dickinson).

While the present invention has been described in terms of the preferred embodiments, it is understood that variations and modifications will occur to those skilled in the art. Therefore, it is intended that the appended claims cover all such equivalent variations that come within the scope of the invention as claimed.

has an activity of the encoded polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23, or is antigenic;

(d) a region of the nucleotide sequence of any of SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 19, or SEQ ID NO: 22, or any of (a) - (c) comprising a fragment of at least about 16 nucleotides;

(e) a nucleotide sequence which hybridizes under moderately or highly stringent conditions to the complement of any of (a) - (d); and

(f) a nucleotide sequence complementary to any of (a) - (d).

3. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one conservative amino acid substitution, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(b) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid insertion, wherein the encoded polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(c) a nucleotide sequence encoding a polypeptide as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ

5. A host cell comprising the vector of Claim 4.

6. The host cell of Claim 5 that is a eukaryotic cell.

7. The host cell of Claim 5 that is a prokaryotic cell.

8. A process of producing an LGR8 polypeptide comprising culturing the host cell of Claim 5 under suitable conditions to express the polypeptide, and optionally isolating the polypeptide from the culture.

9. A polypeptide produced by the process of Claim 8.

10. The process of Claim 8, wherein the nucleic acid molecule comprises promoter DNA other than the promoter DNA for the native LGR8 polypeptide operatively linked to the DNA encoding the LGR8 polypeptide.

11. The isolated nucleic acid molecule according to Claim 2, wherein the percent identity is determined using a computer program selected from the group consisting of GAP, BLASTN, FASTA, BLASTA, BLASTX, BestFit, and the Smith-Waterman algorithm.

12. A process for determining whether a compound inhibits LGR8 polypeptide activity or LGR8 polypeptide production comprising exposing a cell according to any of Claims 5, 6, or 7 to the compound and measuring LGR8 polypeptide activity or LGR8 polypeptide production in said cell.

13. An isolated polypeptide comprising the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

14. An isolated polypeptide comprising the amino acid sequence selected from the group consisting of:

12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one conservative amino acid substitution, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(b) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid insertion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(c) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one amino acid deletion, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23;

(d) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 which has a C- and/or N- terminal truncation, wherein the polypeptide has an activity of the polypeptide set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23; and

(e) the amino acid sequence as set forth in any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23 with at least one modification selected from

22. The selective binding agent of Claim 18 that is a human antibody or fragment thereof.

23. The selective binding agent of Claim 18 that is a polyclonal antibody or fragment thereof.

24. The selective binding agent Claim 18 that is a monoclonal antibody or fragment thereof.

25. The selective binding agent of Claim 18 that is a chimeric antibody or fragment thereof.

26. The selective binding agent of Claim 18 that is a CDR-grafted antibody or fragment thereof.

27. The selective binding agent of Claim 18 that is an antiidiotypic antibody or fragment thereof.

28. The selective binding agent of Claim 18 that is a variable region fragment.

29. The variable region fragment of Claim 28 that is a Fab or a Fab' fragment.

30. A selective binding agent or fragment thereof comprising at least one complementarity determining region with specificity for a polypeptide having the amino acid sequence of any of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 21, or SEQ ID NO: 23.

31. The selective binding agent of Claim 18 that is bound to a detectable label.

41. The polypeptide of Claim 40 that is covalently modified with a water-soluble polymer.

42. The polypeptide of Claim 41, wherein the water-soluble polymer is selected from the group consisting of polyethylene glycol, monomethoxy-polyethylene glycol, dextran, cellulose, poly-(N-vinyl pyrrolidone) polyethylene glycol, propylene glycol homopolymers, polypropylene oxide/ethylene oxide co-polymers, polyoxyethylated polyols, and polyvinyl alcohol.

43. A composition comprising a nucleic acid molecule of any of Claims 1, 2, or 3 and a pharmaceutically acceptable formulation agent.

44. The composition of Claim 43, wherein said nucleic acid molecule is contained in a viral vector.

45. A viral vector comprising a nucleic acid molecule of any of Claims 1, 2, or 3.

46. A fusion polypeptide comprising the polypeptide of any of Claims 13, 14, or 15 fused to a heterologous amino acid sequence.

47. The fusion polypeptide of Claim 46, wherein the heterologous amino acid sequence is an IgG constant domain or fragment thereof.

48. A method for treating, preventing, or ameliorating a medical condition comprising administering to a patient the polypeptide of any of Claims 13, 14, or 15, or the polypeptide encoded by the nucleic acid of any of Claims 1, 2, or 3.

49. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of any of Claims 13, 14, or 15, or the polypeptide encoded by the nucleic acid molecule of any of Claims 1, 2, or 3 in a sample; and

56. A nucleic acid molecule of any of Claims 1, 2, or 3 attached to a solid support.

57. An array of nucleic acid molecules comprising at least one nucleic acid molecule of any of Claims 1, 2, or 3.

58. An isolated polypeptide comprising the amino acid sequence as set forth in SEQ ID NO: 2 with at least one conservative amino acid substitution selected from the group consisting of: isoleucine at position 26; valine at position 41; isoleucine at position 55; aspartic acid at position 78; aspartic acid at position 123; arginine at position 130; valine at position 135; methionine at position 142; leucine at position 166; tyrosine at position 167; lysine at position 201; valine at position 204; isoleucine at position 216; glutamatic acid at position 217; leucine at position 221; leucine at position 240; leucine at position 252; isoleucine at position 277; methionine at position 288; lysine at position 290; isoleucine at position 324; isoleucine at position 341; isoleucine at position 344; aspartic acid at position 350; leucine at position 376; valine at position 420; valine at position 425; valine at position 427; isoleucine at position 434; tyrosine at position 442; arginine at position 444; tyrosine at position 450; isoleucine at position 466; isoleucine at position 471; leucine at position 476; phenylalanine at position 478; glutamatic acid at position 481; histidine at position 485; phenylalanine at position 515; tyrosine at position 521; isoleucine at position 522; tyrosine at position 526; valine at position 531; valine at position 541; isoleucine at position 551; valine at position 552; glutamatic acid at position 561; phenylalanine at position 562; tyrosine at position 566; tyrosine at position 577; aspartic acid at position 579; isoleucine at position 597; isoleucine at position 603; valine at position 616; isoleucine at position 621; isoleucine at position 626; lysine at position 632; leucine at position 649; isoleucine at position 654; valine at position 675; isoleucine at position 682; glutamatic acid at position 700; isoleucine at position 702; tyrosine at position 707; tyrosine at position 709; isoleucine at position 727; valine at position 729; methionine at position 737; methionine at position 745; and leucine at position 749; wherein the polypeptide has an activity of the polypeptide as set forth in SEQ ID NO: 2.

FIG. 1A

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att   48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat   96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
            20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga  144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
        35                  40                  45

tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac  192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
    50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt  240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
65                  70                  75                  80

ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat  288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca  336
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac  384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt  432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac  480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata  528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta tat ctc  576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

aac cac aac tgc atc aca acc ctc aga cct gga ata ttc aaa gac tta  624
Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205
```

FIG. 1B

```
cat cag cta act tgg cta att cta gat gac aat cca ata acc aga att   672
His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg   720
Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

gtt aat aac tac tta gaa gct ctt ccc aag cag atg tgt gcc caa atg   768
Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc   816
Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca gtg ctg ttt ctg   864
Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Phe Leu
        275                 280                 285

cct aga aat caa att ggt ttt gtt cca gag aag aca ttt tct tca tta   912
Pro Arg Asn Gln Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu
    290                 295                 300

aaa aat tta gga gaa ctg gat ctg tct agc aat acg ata acg gag cta   960
Lys Asn Leu Gly Glu Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu
305                 310                 315                 320

tca cct cac ctt ttt aaa gac ttg aag ctt cta caa aag ctg aac ctg  1008
Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asn Leu
                325                 330                 335

tca tcc aat cct ctt atg tat ctt cac aag aac cag ttt gaa agt ctt  1056
Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln Phe Glu Ser Leu
            340                 345                 350

aaa caa ctt cag tct cta gac ctg gaa agg ata gag att cca aat ata  1104
Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile
        355                 360                 365

aac aca cga atg ttt caa ccc atg aag aat ctt tct cac att tat ttc  1152
Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Phe
    370                 375                 380

aaa aac ttt cga tac tgc tcc tat gct ccc cat gtc cga ata tgt atg  1200
Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val Arg Ile Cys Met
385                 390                 395                 400

ccc ttg acg gac ggc att tct tca ttt gag gac ctc ttg gct aac aat  1248
Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn Asn
                405                 410                 415
```

FIG. 1C

```
atc ctc aga ata ttt gtc tgg gtt ata gct ttc att acc tgc ttt gga   1296
Ile Leu Arg Ile Phe Val Trp Val Ile Ala Phe Ile Thr Cys Phe Gly
            420                 425                 430

aat ctt ttt gtc att ggc atg aga tct ttc att aaa gct gaa aat aca   1344
Asn Leu Phe Val Ile Gly Met Arg Ser Phe Ile Lys Ala Glu Asn Thr
        435                 440                 445

act cac gct atg tcc atc aaa atc ctt tgt tgt gct gat tgc ctg atg   1392
Thr His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met
    450                 455                 460

ggt gtt tac ttg ttc ttt gtt ggc att ttc gat ata aaa tac cga ggg   1440
Gly Val Tyr Leu Phe Phe Val Gly Ile Phe Asp Ile Lys Tyr Arg Gly
465                 470                 475                 480

cag tat cag aag tat gcc ttg ctg tgg atg gag agc gtg cag tgc cgc   1488
Gln Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser Val Gln Cys Arg
                485                 490                 495

ctc atg ggg ttc ctg gcc atg ctg tcc acc gaa gtc tct gtt ctg cta   1536
Leu Met Gly Phe Leu Ala Met Leu Ser Thr Glu Val Ser Val Leu Leu
            500                 505                 510

ctg acc tac ttg act ttg gag aag ttc ctg gtc att gtc ttc ccc ttc   1584
Leu Thr Tyr Leu Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe
        515                 520                 525

agt aac att cga cct gga aaa cgg cag acc tca gtc atc ctc att tgc   1632
Ser Asn Ile Arg Pro Gly Lys Arg Gln Thr Ser Val Ile Leu Ile Cys
    530                 535                 540

atc tgg atg gcg gga ttt tta ata gct gta att cca ttt tgg aat aag   1680
Ile Trp Met Ala Gly Phe Leu Ile Ala Val Ile Pro Phe Trp Asn Lys
545                 550                 555                 560

gat tat ttt gga aac ttt tat ggg aaa aat gga gta tgt ttc cca ctt   1728
Asp Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu
                565                 570                 575

tat tat gac caa aca gaa gat att gga agc aaa ggg tat tct ctt gga   1776
Tyr Tyr Asp Gln Thr Glu Asp Ile Gly Ser Lys Gly Tyr Ser Leu Gly
            580                 585                 590

att ttc cta ggt gtg aac ttg ctg gct ttt ctc atc att gtg ttt tcc   1824
Ile Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Ile Ile Val Phe Ser
        595                 600                 605

tat att act atg ttc tgt tcc att caa aaa acc gcc ttg cag acc aca   1872
Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys Thr Ala Leu Gln Thr Thr
    610                 615                 620
```

FIG. 1D

```
gaa gta agg aat tgt ttt gga aga gag gtg gct gtt gca aat cgt ttc   1920
Glu Val Arg Asn Cys Phe Gly Arg Glu Val Ala Val Ala Asn Arg Phe
625                 630                 635                 640

ttt ttt ata gtg ttc tct gat gcc atc tgc tgg att cct gta ttt gta   1968
Phe Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val
                645                 650                 655

gtt aaa atc ctt tcc ctc ttc cgg gtg gaa ata cca gac aca atg act   2016
Val Lys Ile Leu Ser Leu Phe Arg Val Glu Ile Pro Asp Thr Met Thr
            660                 665                 670

tcc tgg ata gtg att ttt ttc ctt cca gtt aac agt gct ttg aat cca   2064
Ser Trp Ile Val Ile Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro
        675                 680                 685

atc ctc tat act ctc aca acc aac ttt ttt aag gac aag ttg aaa cag   2112
Ile Leu Tyr Thr Leu Thr Thr Asn Phe Phe Lys Asp Lys Leu Lys Gln
    690                 695                 700

ctg ctg cac aaa cat cag agg aaa tca att ttc aaa att aaa aaa aaa   2160
Leu Leu His Lys His Gln Arg Lys Ser Ile Phe Lys Ile Lys Lys Lys
705                 710                 715                 720

agt tta tct aca tcc att gtg tgg ata gag gac tcc tct tcc ctg aaa   2208
Ser Leu Ser Thr Ser Ile Val Trp Ile Glu Asp Ser Ser Ser Leu Lys
                725                 730                 735

ctt ggg gtt ttg aac aaa ata aca ctt gga gac agt ata atg aaa cca   2256
Leu Gly Val Leu Asn Lys Ile Thr Leu Gly Asp Ser Ile Met Lys Pro
            740                 745                 750

gtt tcc tag                                                      2265
Val Ser
        755
```

FIG. 2A

```
caa ggt agc atg atc act cct tca tgc caa aaa gga tat ttt ccc tgt   48
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctt acc aag tgc tta ccc cga gct ttt cac tgt gat ggc aag   96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

gga tgg gcg acc ata ttt ggc aca gtg cat gga aat gct aac agc gtg   192
Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

gcc tta aca cag gag tgc ttt cta aaa cag tat cca caa tgc tgt gac   240
Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac tta aag tct gtg   288
Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

ccg atg att tct aac aat gtg aca tta ctg tct ctt aag aaa aac aaa   336
Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac agt ctt cca gat aaa gtt ttc atc aaa tac aca aaa ctt aaa   384
Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

aag ata ttt ctt cag cat aat tgc att aga cac ata tcc agg aaa gca   432
Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

ttt ttt gga tta tgt aat ctg caa ata tta tat ctc aac cac aac tgc   480
Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

atc aca acc ctc aga cct gga ata ttc aaa gac tta cat cag cta act   528
Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

tgg cta att cta gat gac aat cca ata acc aga att tca cag cgc ttg   576
Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg gtt aat aac tac   624
Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205
```

FIG. 2B

```
tta gaa gct ctt ccc aag cag atg tgt gcc caa atg cct caa ctc aac   672
Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc aca aat tct acg   720
Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225                 230                 235                 240

ttt ctg tcg tgc gat tcg ctc aca gtg ctg ttt ctg cct aga aat caa   768
Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
                245                 250                 255

att ggt ttt gtt cca gag aag aca ttt tct tca tta aaa aat tta gga   816
Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
            260                 265                 270

gaa ctg gat ctg tct agc aat acg ata acg gag cta tca cct cac ctt   864
Glu Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
        275                 280                 285

ttt aaa gac ttg aag ctt cta caa aag ctg aac ctg tca tcc aat cct   912
Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
    290                 295                 300

ctt atg tat ctt cac aag aac cag ttt gaa agt ctt aaa caa ctt cag   960
Leu Met Tyr Leu His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln
305                 310                 315                 320

tct cta gac ctg gaa agg ata gag att cca aat ata aac aca cga atg  1008
Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met
                325                 330                 335

ttt caa ccc atg aag aat ctt tct cac att tat ttc aaa aac ttt cga  1056
Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg
            340                 345                 350

tac tgc tcc tat gct ccc cat gtc cga ata tgt atg ccc ttg acg gac  1104
Tyr Cys Ser Tyr Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp
        355                 360                 365

ggc att tct tca ttt gag gac ctc ttg gct aac aat atc ctc aga      1149
Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg
    370                 375                 380
```

FIG. 3A

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att   48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
 1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat   96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
            20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga  144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
        35                  40                  45

tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac  192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
    50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt  240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
65                  70                  75                  80

ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat  288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca  336
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac  384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt  432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac  480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata  528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta tat ctc  576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

aac cac aac tgc atc aca acc ctc aga cct gga ata ttc aaa gac tta  624
Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205
```

FIG. 3B

```
cat cag cta act tgg cta att cta gat gac aat cca ata acc aga att   672
His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg   720
Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

gtt aat aac tac tta gaa gct ctt ccc aag cag atg tgt gcc caa atg   768
Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc   816
Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca gtg ctg gat ctg   864
Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu
        275                 280                 285

tct agc aat acg ata acg gag cta tca cct cac ctt ttt aaa gac ttg   912
Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu
    290                 295                 300

aag ctt cta caa aag ctg aac ctg tca tcc aat cct ctt atg tat ctt   960
Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu
305                 310                 315                 320

cac aag aac cag ttt gaa agt ctt aaa caa ctt cag tct cta gac ctg  1008
His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu
                325                 330                 335

gaa agg ata gag att cca aat ata aac aca cga atg ttt caa ccc atg  1056
Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
            340                 345                 350

aag aat ctt tct cac att tat ttc aaa aac ttt cga tac tgc tcc tat  1104
Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr
        355                 360                 365

gct ccc cat gtc cga ata tgt atg ccc ttg acg gac ggc att tct tca  1152
Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser
    370                 375                 380

ttt gag gac ctc ttg gct aac aat atc ctc aga ata ttt gtc tgg gtt  1200
Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val
385                 390                 395                 400

ata gct ttc att acc tgc ttt gga aat ctt ttt gtc att ggc atg aga  1248
Ile Ala Phe Ile Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg
                405                 410                 415
```

FIG. 3C

```
tct ttc att aaa gct gaa aat aca act cac gct atg tcc atc aaa atc   1296
Ser Phe Ile Lys Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile
            420                 425                 430

ctt tgt tgt gct gat tgc ctg atg ggt gtt tac ttg ttc ttt gtt ggc   1344
Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly
        435                 440                 445

att ttc gat ata aaa tac cga ggg cag tat cag aag tat gcc ttg ctg   1392
Ile Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu
    450                 455                 460

tgg atg gag agc gtg cag tgc cgc ctc atg ggg ttc ctg gcc atg ctg   1440
Trp Met Glu Ser Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu
465                 470                 475                 480

tcc acc gaa gtc tct gtt ctg cta ctg acc tac ttg act ttg gag aag   1488
Ser Thr Glu Val Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys
                485                 490                 495

ttc ctg gtc att gtc ttc ccc ttc agt aac att cga cct gga aaa cgg   1536
Phe Leu Val Ile Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg
            500                 505                 510

cag acc tca gtc atc ctc att tgc atc tgg atg gcg gga ttt tta ata   1584
Gln Thr Ser Val Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile
        515                 520                 525

gct gta att cca ttt tgg aat aag gat tat ttt gga aac ttt tat ggg   1632
Ala Val Ile Pro Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly
    530                 535                 540

aaa aat gga gta tgt ttc cca ctt tat tat gac caa aca gaa gat att   1680
Lys Asn Gly Val Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile
545                 550                 555                 560

gga agc aaa ggg tat tct ctt gga att ttc cta ggt gtg aac ttg ctg   1728
Gly Ser Lys Gly Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu
                565                 570                 575

gct ttt ctc atc att gtg ttt tcc tat att act atg ttc tgt tcc att   1776
Ala Phe Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile
            580                 585                 590

caa aaa acc gcc ttg cag acc aca gaa gta agg aat tgt ttt gga aga   1824
Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg
        595                 600                 605

gag gtg gct gtt gca aat cgt ttc ttt ttt ata gtg ttc tct gat gcc   1872
Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala
    610                 615                 620
```

FIG. 3D

```
atc tgc tgg att cct gta ttt gta gtt aaa atc ctt tcc ctc ttc cgg   1920
Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg
625                 630                 635                 640

gtg gaa ata cca gac aca atg act tcc tgg ata gtg att ttt ttc ctt   1968
Val Glu Ile Pro Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu
                645                 650                 655

cca gtt aac agt gct ttg aat cca atc ctc tat act ctc aca acc aac   2016
Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn
            660                 665                 670

ttt ttt aag gac aag ttg aaa cag ctg ctg cac aaa cat cag agg aaa   2064
Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys
        675                 680                 685

tca att ttc aaa att aaa aaa aaa agt tta tct aca tcc att gtg tgg   2112
Ser Ile Phe Lys Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp
    690                 695                 700

ata gag gac tcc tct tcc ctg aaa ctt ggg gtt ttg aac aaa ata aca   2160
Ile Glu Asp Ser Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr
705                 710                 715                 720

ctt gga gac agt ata atg aaa cca gtt tcc tag                      2193
Leu Gly Asp Ser Ile Met Lys Pro Val Ser
                725                 730
```

FIG. 4A

```
caa ggt agc atg atc act cct tca tgc caa aaa gga tat ttt ccc tgt   48
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctt acc aag tgc tta ccc cga gct ttt cac tgt gat ggc aag   96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

gga tgg gcg acc ata ttt ggc aca gtg cat gga aat gct aac agc gtg   192
Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

gcc tta aca cag gag tgc ttt cta aaa cag tat cca caa tgc tgt gac   240
Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac tta aag tct gtg   288
Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

ccg atg att tct aac aat gtg aca tta ctg tct ctt aag aaa aac aaa   336
Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac agt ctt cca gat aaa gtt ttc atc aaa tac aca aaa ctt aaa   384
Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

aag ata ttt ctt cag cat aat tgc att aga cac ata tcc agg aaa gca   432
Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

ttt ttt gga tta tgt aat ctg caa ata tta tat ctc aac cac aac tgc   480
Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

atc aca acc ctc aga cct gga ata ttc aaa gac tta cat cag cta act   528
Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

tgg cta att cta gat gac aat cca ata acc aga att tca cag cgc ttg   576
Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg gtt aat aac tac   624
Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205
```

FIG. 4B

```
tta gaa gct ctt ccc aag cag atg tgt gcc caa atg cct caa ctc aac   672
Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc aca aat tct acg   720
Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225                 230                 235                 240

ttt ctg tcg tgc gat tcg ctc aca gtg ctg gat ctg tct agc aat acg   768
Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr
                245                 250                 255

ata acg gag cta tca cct cac ctt ttt aaa gac ttg aag ctt cta caa   816
Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln
            260                 265                 270

aag ctg aac ctg tca tcc aat cct ctt atg tat ctt cac aag aac cag   864
Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln
        275                 280                 285

ttt gaa agt ctt aaa caa ctt cag tct cta gac ctg gaa agg ata gag   912
Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu
    290                 295                 300

att cca aat ata aac aca cga atg ttt caa ccc atg aag aat ctt tct   960
Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
305                 310                 315                 320

cac att tat ttc aaa aac ttt cga tac tgc tcc tat gct ccc cat gtc   1008
His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
                325                 330                 335

cga ata tgt atg ccc ttg acg gac ggc att tct tca ttt gag gac ctc   1056
Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
            340                 345                 350

ttg gct aac aat atc ctc aga                                      1077
Leu Ala Asn Asn Ile Leu Arg
        355
```

FIG. 5A

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att    48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
 1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat    96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
            20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga   144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
        35                  40                  45

tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac   192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
    50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt   240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
65                  70                  75                  80

ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat   288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca   336
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac   384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt   432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac   480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata   528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta att cta   576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Ile Leu
            180                 185                 190

gat gac aat cca ata acc aga att tca cag cgc ttg ttt acg gga tta   624
Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu Phe Thr Gly Leu
        195                 200                 205
```

FIG. 5B

```
aat tcc ttg ttt ttc ctg tct atg gtt aat aac tac tta gaa gct ctt   672
Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr Leu Glu Ala Leu
    210                 215                 220

ccc aag cag atg tgt gcc caa atg cct caa ctc aac tgg gtg gat ttg   720
Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn Trp Val Asp Leu
225                 230                 235                 240

gaa ggc aat aga ata aag tat ctc aca aat tct acg ttt ctg tcg tgc   768
Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr Phe Leu Ser Cys
                245                 250                 255

gat tcg ctc aca gtg ctg gat ctg tct agc aat acg ata acg gag cta   816
Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu
            260                 265                 270

tca cct cac ctt ttt aaa gac ttg aag ctt cta caa aag cta gac ctg   864
Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asp Leu
        275                 280                 285

gaa agg ata gag att cca aat ata aac aca cga atg ttt caa ccc atg   912
Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
    290                 295                 300

aag aat ctt tct cac att tat ttc aaa aac ttt cga tac tgc tcc tat   960
Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr
305                 310                 315                 320

gct ccc cat gtc cga ata tgt atg ccc ttg acg gac ggc att tct tca  1008
Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser
                325                 330                 335

ttt gag gac ctc ttg gct aac aat atc ctc aga ata ttt gtc tgg gtt  1056
Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val
            340                 345                 350

ata gct ttc att acc tgc ttt gga aat ctt ttt gtc att ggc atg aga  1104
Ile Ala Phe Ile Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg
        355                 360                 365

tct ttc att aaa gct gaa aat aca act cac gct atg tcc atc aaa atc  1152
Ser Phe Ile Lys Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile
    370                 375                 380

ctt tgt tgt gct gat tgc ctg atg ggt gtt tac ttg ttc ttt gtt ggc  1200
Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly
385                 390                 395                 400

att ttc gat ata aaa tac cga ggg cag tat cag aag tat gcc ttg ctg  1248
Ile Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu
                405                 410                 415
```

FIG. 5C

```
tgg atg gag agc gtg cag tgc cgc ctc atg ggg ttc ctg gcc atg ctg   1296
Trp Met Glu Ser Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu
            420                 425                 430

tcc acc gaa gtc tct gtt ctg cta ctg acc tac ttg act ttg gag aag   1344
Ser Thr Glu Val Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys
        435                 440                 445

ttc ctg gtc att gtc ttc ccc ttc agt aac att cga cct gga aaa cgg   1392
Phe Leu Val Ile Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg
    450                 455                 460

cag acc tca gtc atc ctc att tgc atc tgg atg gcg gga ttt tta ata   1440
Gln Thr Ser Val Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile
465                 470                 475                 480

gct gta att cca ttt tgg aat aag gat tat ttt gga aac ttt tat ggg   1488
Ala Val Ile Pro Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly
                485                 490                 495

aaa aat gga gta tgt ttc cca ctt tat tat gac caa aca gaa gat att   1536
Lys Asn Gly Val Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile
            500                 505                 510

gga agc aaa ggg tat tct ctt gga att ttc cta ggt gtg aac ttg ctg   1584
Gly Ser Lys Gly Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu
        515                 520                 525

gct ttt ctc atc att gtg ttt tcc tat att act atg ttc tgt tcc att   1632
Ala Phe Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile
    530                 535                 540

caa aaa acc gcc ttg cag acc aca gaa gta agg aat tgt ttt gga aga   1680
Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg
545                 550                 555                 560

gag gtg gct gtt gca aat cgt ttc ttt ttt ata gtg ttc tct gat gcc   1728
Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala
                565                 570                 575

atc tgc tgg att cct gta ttt gta gtt aaa atc ctt tcc ctc ttc cgg   1776
Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg
            580                 585                 590

gtg gaa ata cca gac aca atg act tcc tgg ata gtg att ttt ttc ctt   1824
Val Glu Ile Pro Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu
        595                 600                 605

cca gtt aac agt gct ttg aat cca atc ctc tat act ctc aca acc aac   1872
Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn
    610                 615                 620
```

FIG. 5D

```
ttt ttt aag gac aag ttg aaa cag ctg ctg cac aaa cat cag agg aaa   1920
Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys
625                 630                 635                 640

tca att ttc aaa att aaa aaa aaa agt tta tct aca tcc att gtg tgg   1968
Ser Ile Phe Lys Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp
                645                 650                 655

ata gag gac tcc tct tcc ctg aaa ctt ggg gtt ttg aac aaa ata aca   2016
Ile Glu Asp Ser Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr
            660                 665                 670

ctt gga gac agt ata atg aaa cca gtt tcc tag                      2049
Leu Gly Asp Ser Ile Met Lys Pro Val Ser
        675                 680
```

FIG. 6A

```
caa ggt agc atg atc act cct tca tgc caa aaa gga tat ttt ccc tgt   48
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctt acc aag tgc tta ccc cga gct ttt cac tgt gat ggc aag   96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

gga tgg gcg acc ata ttt ggc aca gtg cat gga aat gct aac agc gtg   192
Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

gcc tta aca cag gag tgc ttt cta aaa cag tat cca caa tgc tgt gac   240
Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac tta aag tct gtg   288
Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

ccg atg att tct aac aat gtg aca tta ctg tct ctt aag aaa aac aaa   336
Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac agt ctt cca gat aaa gtt ttc atc aaa tac aca aaa ctt aaa   384
Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

aag ata ttt ctt cag cat aat tgc att aga cac ata tcc agg aaa gca   432
Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

ttt ttt gga tta tgt aat ctg caa ata tta att cta gat gac aat cca   480
Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Ile Leu Asp Asp Asn Pro
145                 150                 155                 160

ata acc aga att tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt   528
Ile Thr Arg Ile Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe
                165                 170                 175

ttc ctg tct atg gtt aat aac tac tta gaa gct ctt ccc aag cag atg   576
Phe Leu Ser Met Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met
            180                 185                 190

tgt gcc caa atg cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga   624
Cys Ala Gln Met Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg
        195                 200                 205
```

FIG. 6B

```
ata aag tat ctc aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca   672
Ile Lys Tyr Leu Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr
    210                 215                 220

gtg ctg gat ctg tct agc aat acg ata acg gag cta tca cct cac ctt   720
Val Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
225                 230                 235                 240

ttt aaa gac ttg aag ctt cta caa aag cta gac ctg gaa agg ata gag   768
Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asp Leu Glu Arg Ile Glu
                245                 250                 255

att cca aat ata aac aca cga atg ttt caa ccc atg aag aat ctt tct   816
Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
            260                 265                 270

cac att tat ttc aaa aac ttt cga tac tgc tcc tat gct ccc cat gtc   864
His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
        275                 280                 285

cga ata tgt atg ccc ttg acg gac ggc att tct tca ttt gag gac ctc   912
Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
    290                 295                 300

ttg gct aac aat atc ctc aga                                      933
Leu Ala Asn Asn Ile Leu Arg
305                 310
```

FIG. 7A

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att   48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat   96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
            20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga  144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
        35                  40                  45

tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac  192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
    50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt  240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
65                  70                  75                  80

ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat  288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca  336
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac  384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt  432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac  480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata  528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta tat ctc  576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

aac cac aac tgc atc aca acc ctc aga cct gga ata ttc aaa gac tta  624
Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205
```

FIG. 7B

```
cat cag cta act tgg cta att cta gat gac aat cca ata acc aga att   672
His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg   720
Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

gtt aat aac tac tta gaa gct ctt ccc aag cag atg tgt gcc caa atg   768
Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc   816
Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca gtg ctg gat ctg   864
Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu
        275                 280                 285

tct agc aat acg ata acg gag cta tca cct cac ctt ttt aaa gac ttg   912
Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu
    290                 295                 300

aag ctt cta caa aag ctg aac ctg tca tcc aat cct ctt atg tat ctt   960
Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu
305                 310                 315                 320

cac aag aac cag ttt gaa agt ctt aaa caa ctt cag tct cta gac ctg  1008
His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu
                325                 330                 335

gaa agg ata gag att cca aat ata aac aca cga atg ttt caa ccc atg  1056
Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
            340                 345                 350

aag aat ctt tct cac ata gtt caa tat tat gat gtg ccg aca tga      1101
Lys Asn Leu Ser His Ile Val Gln Tyr Tyr Asp Val Pro Thr
        355                 360                 365
```

FIG. 8A

```
atg tgg ctc cta ctt cat gtc atc ctt ctg aca gag gtc aaa gat ttt   48
Met Trp Leu Leu Leu His Val Ile Leu Leu Thr Glu Val Lys Asp Phe
 1               5                  10                  15

gca ctg gct gac agc agt atg gtg gct cct ctg tgc ccc aaa ggg tat   96
Ala Leu Ala Asp Ser Ser Met Val Ala Pro Leu Cys Pro Lys Gly Tyr
            20                  25                  30

ttt ccc tgt ggg aat ctc acc aaa tgc ttg ccc cga gcc ttt cac tgc  144
Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys
        35                  40                  45

gat ggt gtg gat gat tgc ggg aat ggt gcc gac gag gac aac tgt ggt  192
Asp Gly Val Asp Asp Cys Gly Asn Gly Ala Asp Glu Asp Asn Cys Gly
    50                  55                  60

gac act agt gga tgg aca acc ata ttt ggc aca gtc cat ggg aat gtc  240
Asp Thr Ser Gly Trp Thr Thr Ile Phe Gly Thr Val His Gly Asn Val
65                  70                  75                  80

aat aaa gtg aca ttg aca cag gag tgc ttt ctc agc cag tat cca cag  288
Asn Lys Val Thr Leu Thr Gln Glu Cys Phe Leu Ser Gln Tyr Pro Gln
                85                  90                  95

cac tgt tac tgc aga gaa aat gaa ctg gaa tgt gta aag gct gac tta  336
His Cys Tyr Cys Arg Glu Asn Glu Leu Glu Cys Val Lys Ala Asp Leu
            100                 105                 110

aaa gct gtg cca aag gtt tcc agc aac gta aca tta cta tct ctt aag  384
Lys Ala Val Pro Lys Val Ser Ser Asn Val Thr Leu Leu Ser Leu Lys
        115                 120                 125

aaa aac aaa atc cac aga ctt cca gtc aag gtc ttc agc aga tac aca  432
Lys Asn Lys Ile His Arg Leu Pro Val Lys Val Phe Ser Arg Tyr Thr
    130                 135                 140

gaa ctc aga aag ata tac ctt cag cac aac tgc atc aca cac atc tcc  480
Glu Leu Arg Lys Ile Tyr Leu Gln His Asn Cys Ile Thr His Ile Ser
145                 150                 155                 160

agg aga gca ttc ctt gga tta cat aat cta caa ata ctg tat ctc agc  528
Arg Arg Ala Phe Leu Gly Leu His Asn Leu Gln Ile Leu Tyr Leu Ser
                165                 170                 175

cat aac tgc att acc tct ctc agg cct ggg ata ttc aaa gac ttg cat  576
His Asn Cys Ile Thr Ser Leu Arg Pro Gly Ile Phe Lys Asp Leu His
            180                 185                 190

cag ctt gct tgg cta att tta gat gac aac ccg atc acc aga atc tca  624
Gln Leu Ala Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser
        195                 200                 205
```

FIG. 8B

```
cag aag tcc ttt atg ggg tta aac tcc ttg ttt ttc ttg tcc atg gtg   672
Gln Lys Ser Phe Met Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val
    210                 215                 220

ggt aac cgg ctc gag gcc ctt cct gaa aca ttg tgt gct cag atg cct   720
Gly Asn Arg Leu Glu Ala Leu Pro Glu Thr Leu Cys Ala Gln Met Pro
225                 230                 235                 240

caa ctc aac tgg gtg gat ctg gca aac aat gga ata aag tac ata acg   768
Gln Leu Asn Trp Val Asp Leu Ala Asn Asn Gly Ile Lys Tyr Ile Thr
                245                 250                 255

aac tcc acc ttc cta acg tgc gac tcg ctc acg gtt ctg ttt ctg cct   816
Asn Ser Thr Phe Leu Thr Cys Asp Ser Leu Thr Val Leu Phe Leu Pro
            260                 265                 270

aga aat caa att ggt ttt gtt cca gag aag aca ttt tct tca tta aaa   864
Arg Asn Gln Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys
        275                 280                 285

aat tta gga gaa ctg gac ctg tct agc aat atg ata aca aaa ctc cca   912
Asn Leu Gly Glu Leu Asp Leu Ser Ser Asn Met Ile Thr Lys Leu Pro
    290                 295                 300

gtc cac ctt ttc agc gac ctt cat ctt ctc cag aag ctg aac ctg tca   960
Val His Leu Phe Ser Asp Leu His Leu Leu Gln Lys Leu Asn Leu Ser
305                 310                 315                 320

tcc aac cct ctt ctg tat gtc cac aag aac cag ttt gga agt ctc aaa  1008
Ser Asn Pro Leu Leu Tyr Val His Lys Asn Gln Phe Gly Ser Leu Lys
                325                 330                 335

caa ctt cag tct cta gac ctg gaa agg ata gag att cca aac ata agc  1056
Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Ser
            340                 345                 350

aca gga atg ttc cag cca atg aag aac ctt tct cac att tat ttg aaa  1104
Thr Gly Met Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Leu Lys
        355                 360                 365

acc ttt cga tac tgc tcc tat gtc ccc cat gtc cga atc tgt atg ccg  1152
Thr Phe Arg Tyr Cys Ser Tyr Val Pro His Val Arg Ile Cys Met Pro
    370                 375                 380

tcg act gat ggt att tct tcg tct gag gac ctc ttg gct aac ggt atc  1200
Ser Thr Asp Gly Ile Ser Ser Ser Glu Asp Leu Leu Ala Asn Gly Ile
385                 390                 395                 400

ctc aga gtg tct gtc tgg gtt ata gct ttc att acc tgc gtt ggg aat  1248
Leu Arg Val Ser Val Trp Val Ile Ala Phe Ile Thr Cys Val Gly Asn
                405                 410                 415
```

FIG. 8C

```
ttc ctt gtc ata gcc gtg aga tct ctc att aag gct gag aat aca act   1296
Phe Leu Val Ile Ala Val Arg Ser Leu Ile Lys Ala Glu Asn Thr Thr
            420                 425                 430

cac gct atg tcc atc aaa atc ctt tgt tgt gcc gat tgc ctg atg ggg   1344
His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met Gly
        435                 440                 445

gtg tac ctg ttc tcc gtg ggc gtc ttt gac atc aag tac cga ggg cag   1392
Val Tyr Leu Phe Ser Val Gly Val Phe Asp Ile Lys Tyr Arg Gly Gln
    450                 455                 460

tat cag aag tat gcg ctg ctg tgg atg gag agt gtg ccc tgc cgc ctg   1440
Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser Val Pro Cys Arg Leu
465                 470                 475                 480

ctg ggc ttc ctg gcc acg ctg tcc aca gag gtc tcg gtg ctg ctg ctg   1488
Leu Gly Phe Leu Ala Thr Leu Ser Thr Glu Val Ser Val Leu Leu Leu
                485                 490                 495

aca ttc ctg acg ctg gag aag ttc ctt gtc ata gta ttc cct ttc agc   1536
Thr Phe Leu Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe Ser
            500                 505                 510

aac ctg cgc ctg ggc aag cgc cag act gct gtg gcc ctc gcc agc atc   1584
Asn Leu Arg Leu Gly Lys Arg Gln Thr Ala Val Ala Leu Ala Ser Ile
        515                 520                 525

tgg gtg gtg gga ttt ctc ata gcg gcc gtt ccg ttc acc aga gag gat   1632
Trp Val Val Gly Phe Leu Ile Ala Ala Val Pro Phe Thr Arg Glu Asp
    530                 535                 540

tat ttc ggc aac ttt tat ggg aaa aat gga gtc tgc ttc cca ctt cat   1680
Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu His
545                 550                 555                 560

tat gac caa gca gaa gat ttt gga agt aga ggg tac tcc ctt ggg att   1728
Tyr Asp Gln Ala Glu Asp Phe Gly Ser Arg Gly Tyr Ser Leu Gly Ile
                565                 570                 575

ttc cta ggt gtg aac ttg ctg gct ttc ctc gtc atc gtg att tcc tat   1776
Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Val Ile Val Ile Ser Tyr
            580                 585                 590

gtc acc atg ttc tgc tcc att cat aaa aca gcc ctt cag act gca gaa   1824
Val Thr Met Phe Cys Ser Ile His Lys Thr Ala Leu Gln Thr Ala Glu
        595                 600                 605

gtg agg agc cac atc ggg aag gag gtg gct gtt gca aac cgg ttc ttt   1872
Val Arg Ser His Ile Gly Lys Glu Val Ala Val Ala Asn Arg Phe Phe
    610                 615                 620
```

FIG. 8D

```
ttt atc gtg ttc tct gat gcc atc tgc tgg atc cct gtg ttt gtc gtt   1920
Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val Val
625                 630                 635                 640

aag atc ctg tct ctc ctt caa gtg gag ata cca ggc aca atc act tcc   1968
Lys Ile Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr Ile Thr Ser
                645                 650                 655

tgg atc gtg gtt ttt ttc ctt ccg gtg aac agc gcc tta aac ccc atc   2016
Trp Ile Val Val Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro Ile
            660                 665                 670

ctc tac act ctg acg acc tcc ttt ttt aag gac aag ttg aaa cag ttg   2064
Leu Tyr Thr Leu Thr Thr Ser Phe Phe Lys Asp Lys Leu Lys Gln Leu
        675                 680                 685

ctg cac aaa cat cgg agg aaa ccc atc ttc aaa gtt aag aag aaa agt   2112
Leu His Lys His Arg Arg Lys Pro Ile Phe Lys Val Lys Lys Lys Ser
    690                 695                 700

tta tcc gct tcc att gtg tgg aca gac gag tct tca ctt aaa ctt gga   2160
Leu Ser Ala Ser Ile Val Trp Thr Asp Glu Ser Ser Leu Lys Leu Gly
705                 710                 715                 720

gtg ttg agc aaa ata gcc ctt ggg gac agt ata atg aag ccg gtc tcc   2208
Val Leu Ser Lys Ile Ala Leu Gly Asp Ser Ile Met Lys Pro Val Ser
                725                 730                 735

ccg tag                                                          2214
Pro
```

FIG. 9A

```
gac agc agt atg gtg gct cct ctg tgc ccc aaa ggg tat ttt ccc tgt   48
Asp Ser Ser Met Val Ala Pro Leu Cys Pro Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctc acc aaa tgc ttg ccc cga gcc ttt cac tgc gat ggt gtg   96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Val
            20                  25                  30

gat gat tgc ggg aat ggt gcc gac gag gac aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Asp Asn Cys Gly Asp Thr Ser
        35                  40                  45

gga tgg aca acc ata ttt ggc aca gtc cat ggg aat gtc aat aaa gtg   192
Gly Trp Thr Thr Ile Phe Gly Thr Val His Gly Asn Val Asn Lys Val
    50                  55                  60

aca ttg aca cag gag tgc ttt ctc agc cag tat cca cag cac tgt tac   240
Thr Leu Thr Gln Glu Cys Phe Leu Ser Gln Tyr Pro Gln His Cys Tyr
65                  70                  75                  80

tgc aga gaa aat gaa ctg gaa tgt gta aag gct gac tta aaa gct gtg   288
Cys Arg Glu Asn Glu Leu Glu Cys Val Lys Ala Asp Leu Lys Ala Val
                85                  90                  95

cca aag gtt tcc agc aac gta aca tta cta tct ctt aag aaa aac aaa   336
Pro Lys Val Ser Ser Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac aga ctt cca gtc aag gtc ttc agc aga tac aca gaa ctc aga   384
Ile His Arg Leu Pro Val Lys Val Phe Ser Arg Tyr Thr Glu Leu Arg
        115                 120                 125

aag ata tac ctt cag cac aac tgc atc aca cac atc tcc agg aga gca   432
Lys Ile Tyr Leu Gln His Asn Cys Ile Thr His Ile Ser Arg Arg Ala
    130                 135                 140

ttc ctt gga tta cat aat cta caa ata ctg tat ctc agc cat aac tgc   480
Phe Leu Gly Leu His Asn Leu Gln Ile Leu Tyr Leu Ser His Asn Cys
145                 150                 155                 160

att acc tct ctc agg cct ggg ata ttc aaa gac ttg cat cag ctt gct   528
Ile Thr Ser Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Ala
                165                 170                 175

tgg cta att tta gat gac aac ccg atc acc aga atc tca cag aag tcc   576
Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Lys Ser
            180                 185                 190

ttt atg ggg tta aac tcc ttg ttt ttc ttg tcc atg gtg ggt aac cgg   624
Phe Met Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Gly Asn Arg
        195                 200                 205
```

FIG. 9B

```
ctc gag gcc ctt cct gaa aca ttg tgt gct cag atg cct caa ctc aac   672
Leu Glu Ala Leu Pro Glu Thr Leu Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

tgg gtg gat ctg gca aac aat gga ata aag tac ata acg aac tcc acc   720
Trp Val Asp Leu Ala Asn Asn Gly Ile Lys Tyr Ile Thr Asn Ser Thr
225                 230                 235                 240

ttc cta acg tgc gac tcg ctc acg gtt ctg ttt ctg cct aga aat caa   768
Phe Leu Thr Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
                245                 250                 255

att ggt ttt gtt cca gag aag aca ttt tct tca tta aaa aat tta gga   816
Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
            260                 265                 270

gaa ctg gac ctg tct agc aat atg ata aca aaa ctc cca gtc cac ctt   864
Glu Leu Asp Leu Ser Ser Asn Met Ile Thr Lys Leu Pro Val His Leu
        275                 280                 285

ttc agc gac ctt cat ctt ctc cag aag ctg aac ctg tca tcc aac cct   912
Phe Ser Asp Leu His Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
    290                 295                 300

ctt ctg tat gtc cac aag aac cag ttt gga agt ctc aaa caa ctt cag   960
Leu Leu Tyr Val His Lys Asn Gln Phe Gly Ser Leu Lys Gln Leu Gln
305                 310                 315                 320

tct cta gac ctg gaa agg ata gag att cca aac ata agc aca gga atg   1008
Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Ser Thr Gly Met
                325                 330                 335

ttc cag cca atg aag aac ctt tct cac att tat ttg aaa acc ttt cga   1056
Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Leu Lys Thr Phe Arg
            340                 345                 350

tac tgc tcc tat gtc ccc cat gtc cga atc tgt atg ccg tcg act gat   1104
Tyr Cys Ser Tyr Val Pro His Val Arg Ile Cys Met Pro Ser Thr Asp
        355                 360                 365

ggt att tct tcg tct gag gac ctc ttg gct aac ggt                  1140
Gly Ile Ser Ser Ser Glu Asp Leu Leu Ala Asn Gly
    370                 375                 380
```

FIG. 10A

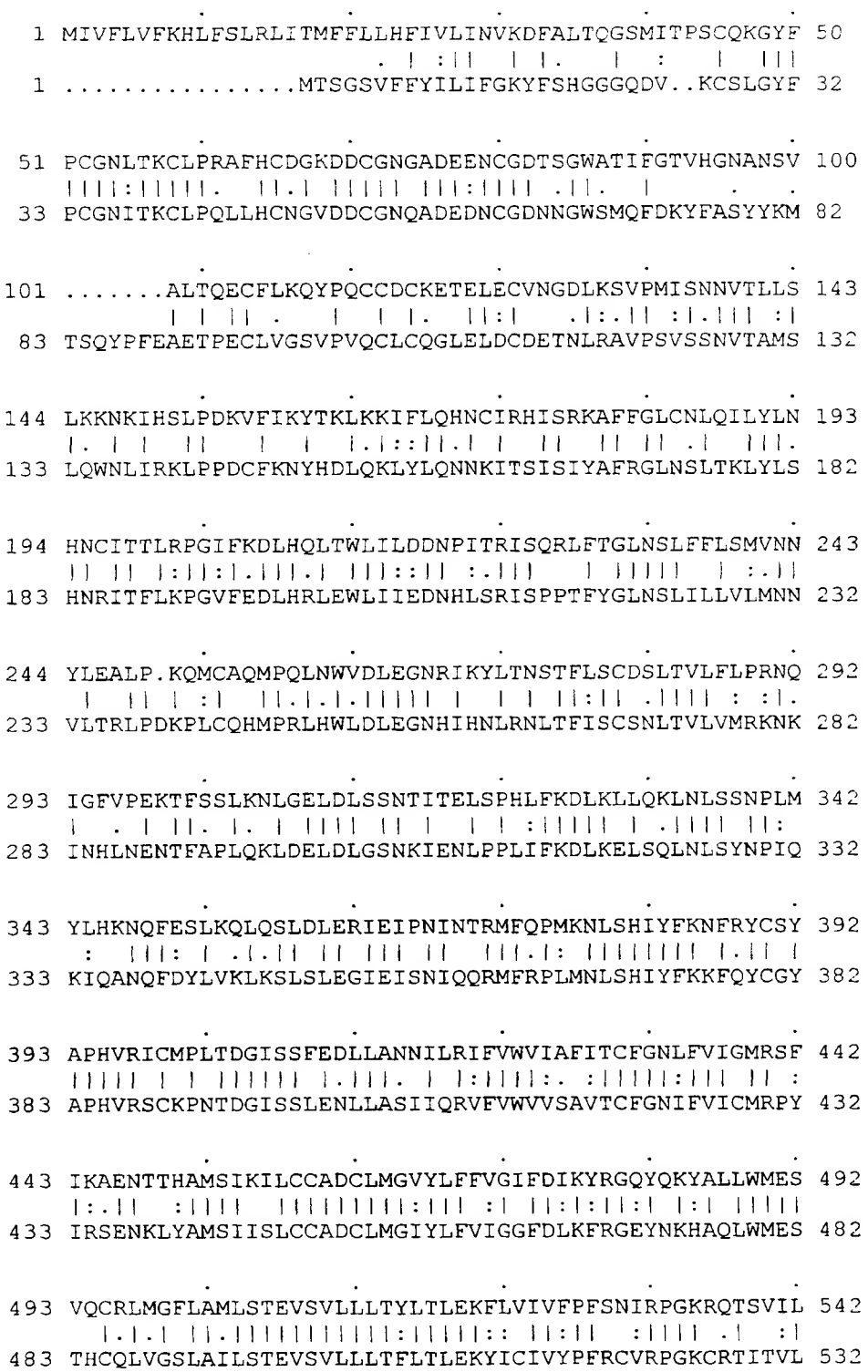

FIG. 10B

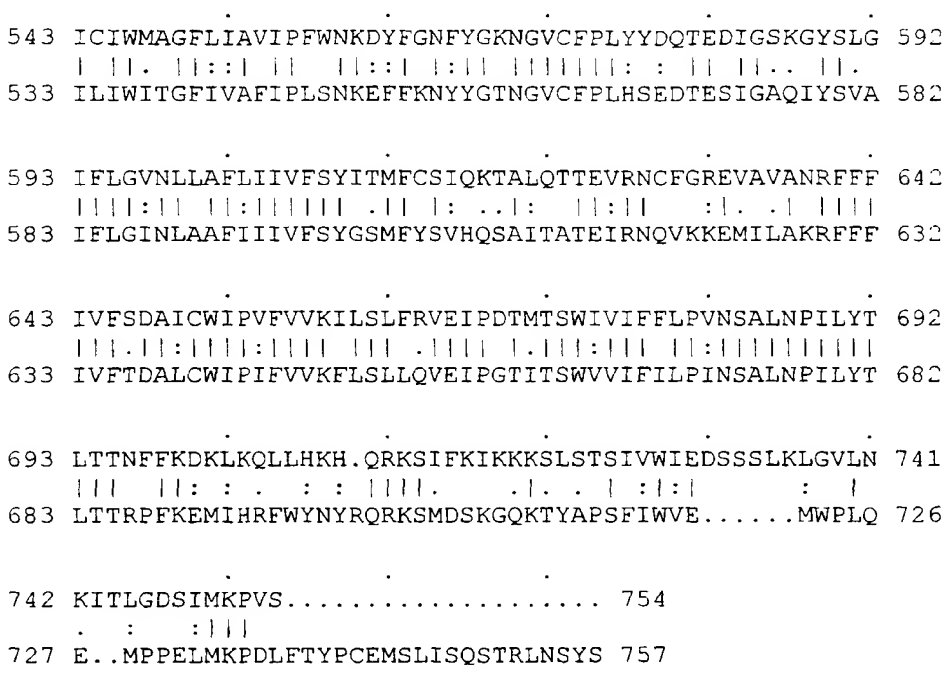
543 ICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQTEDIGSKGYSLG 592
533 ILIWITGFIVAFIPLSNKEFFKNYYGTNGVCFPLHSEDTESIGAQIYSVA 582
593 IFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFF 642
583 IFLGINLAAFIIIVFSYGSMFYSVHQSAITATEIRNQVKKEMILAKRFFF 632
643 IVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSALNPILYT 692
633 IVFTDALCWIPIFVVKFLSLLQVEIPGTITSWVVIFILPINSALNPILYT 682
693 LTTNFFKDKLKQLLHKH.QRKSIFKIKKKSLSTSIVWIEDSSSLKLGVLN 741
683 LTTRPFKEMIHRFWYNYRQRKSMDSKGQKTYAPSFIWVE......MWPLQ 726
742 KITLGDSIMKPVS..... 754
727 E..MPPELMKPDLFTYPCEMSLISQSTRLNSYS 757

FIG. 11A


1 QGSMITPSCQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGW 50
1 DSSMVAPLCPKGYFPCGNLTKCLPRAFHCDGVDDCGNGADEDNCGDTSGW 50
51 ATIFGTVHGNANSVALTQECFLKQYPQCCDCKETELECVNGDLKSVPMIS 100
51 TTIFGTVHGNVNKVTLTQECFLSQYPQHCYCRENELECVKADLKAVPKVS 100
101 NNVTLLSLKKNKIHSLPDKVFIKYTKLKKIFLQHNCIRHISRKAFFGLCN 150
101 SNVTLLSLKKNKIHRLPVKVFSRYTELRKIYLQHNCITHISRRAFLGLHN 150
151 LQILYLNHNCITTLRPGIFKDLHQLTWLILDDNPITRISQRLFTGLNSLF 200
151 LQILYLSHNCITSLRPGIFKDLHQLAWLILDDNPITRISQKSFMGLNSLF 200
201 FLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNSTFLSCDSLTVL 250
201 FLSMVGNRLEALPETLCAQMPQLNWVDLANNGIKYITNSTFLTCDSLTVL 250
251 FLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKLNL 300
251 FLPRNQIGFVPEKTFSSLKNLGELDLSSNMITKLPVHLFSDLHLLQKLNL 300
301 SSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKN 350
301 SSNPLLYVHKNQFGSLKQLQSLDLERIEIPNISTGMFQPMKNLSHIYLKT 350
351 FRYCSYAPHVRICMPLTDGISSFEDLLANNILRIFVWVIAFITCFGNLFV 400
351 FRYCSYVPHVRICMPSTDGISSSEDLLANGILRVSVWVIAFITCVGNFLV 400
401 IGMRSFIKAENTTHAMSIKILCCADCLMGVYLFFVGIFDIKYRGQYQKYA 450
401 IAVRSLIKAENTTHAMSIKILCCADCLMGVYLFSVGVFDIKYRGQYQKYA 450
451 LLWMESVQCRLMGFLAMLSTEVSVLLLTYLTLEKFLVIVFPFSNIRPGKR 500
451 LLWMESVPCRLLGFLATLSTEVSVLLLTFLTLEKFLVIVFPFSNLRLGKR 500
501 QTSVILICIWMAGFLIAVIPFWNKDYFGNFYGKNGVCFPLYYDQTEDIGS 550
501 QTAVALASIWVVGFLIAAVPFTREDYFGNFYGKNGVCFPLHYDQAEDFGS 550

FIG. 11B

```
                .         .         .         .         .
551 KGYSLGIFLGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAV 600
    :|||||||||||||||||:|| ||:|||||| |||||| |||.  |:||||
551 RGYSLGIFLGVNLLAFLVIVISYVTMFCSIHKTALQTAEVRSHIGKEVAV 600

                .         .         .         .         .
601 ANRFFFIVFSDAICWIPVFVVKILSLFRVEIPDTMTSWIVIFFLPVNSAL 650
    |||||||||||||||||||||||||| .|||| |.|||||:|||||||||
601 ANRFFFIVFSDAICWIPVFVVKILSLLQVEIPGTITSWIVVFFLPVNSAL 650

                .         .         .         .         .
651 NPILYTLTTNFFKDKLKQLLHKHQRKSIFKIKKKSLSTSIVWIEDSSSLK 700
    |||||||||.|||||||||||||.|| |||:|||||| |||| | ||||
651 NPILYTLTTSFFKDKLKQLLHKHRRKPIFKVKKKSLSASIVW.TDESSLK 699

                .         .
701 LGVLNKITLGDSIMKPVS.. 718
    ||||.|| ||||||||||
700 LGVLSKIALGDSIMKPVSP  718
```

FIG. 12


1 QGSMITPSCQKGYFPCGNLTKCLPRAFHCDGKDDCGNGADEENCGDTSGW 50
1 DSSMVAPLCPKGYFPCGNLTKCLPRAFHCDGVDDCGNGADEDNCGDTSGW 50
51 ATIFGTVHGNANSVALTQECFLKQYPQCCDCKETELECVNGDLKSVPMIS 100
51 TTIFGTVHGNVNKVTLTQECFLSQYPQHCYCRENELECVKADLKAVPKVS 100
101 NNVTLLSLKKNKIHSLPDKVFIKYTKLKKIFLQHNCIRHISRKAFFGLCN 150
101 SNVTLLSLKKNKIHRLPVKVFSRYTELRKIYLQHNCITHISRRAFLGLHN 150
151 LQILYLNHNCITTLRPGIFKDLHQLTWLILDDNPITRISQRLFTGLNSLF 200
151 LQILYLSHNCITSLRPGIFKDLHQLAWLILDDNPITRISQKSFMGLNSLF 200
201 FLSMVNNYLEALPKQMCAQMPQLNWVDLEGNRIKYLTNSTFLSCDSLTVL 250
201 FLSMVGNRLEALPETLCAQMPQLNWVDLANNGIKYITNSTFLTCDSLTVL 250
251 FLPRNQIGFVPEKTFSSLKNLGELDLSSNTITELSPHLFKDLKLLQKLNL 300
251 FLPRNQIGFVPEKTFSSLKNLGELDLSSNMITKLPVHLFSDLHLLQKLNL 300
301 SSNPLMYLHKNQFESLKQLQSLDLERIEIPNINTRMFQPMKNLSHIYFKN 350
301 SSNPLLYVHKNQFGSLKQLQSLDLERIEIPNISTGMFQPMKNLSHIYLKT 350
351 FRYCSYAPHVRICMPLTDGISSFEDLLANNILR 383
351 FRYCSYVPHVRICMPSTDGISSSEDLLANG... 380

SEQUENCE LISTING

<110> Paszty, Christopher J.
Gong, Jianhua
Daugherty, Betsy
Rogers, Norma

<120> Leucine-Rich G Protein Coupled Receptor-8 Molecules and Uses Thereof

<130> 00-1229-A

<140>
<141>

<150> 60/224,455
<151> 2000-08-10

<160> 42

<170> PatentIn Ver. 2.0

<210> 1
<211> 2265
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(2262)

<220>
<221> sig_peptide
<222> (1)..(108)

<400> 1

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att   48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat   96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
             20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga   144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
         35                  40                  45

tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac   192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
     50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt   240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
 65                  70                  75                  80

ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat   288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                 85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca   336
```

```
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac   384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt   432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac   480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata   528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta tat ctc   576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

aac cac aac tgc atc aca acc ctc aga cct gga ata ttc aaa gac tta   624
Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205

cat cag cta act tgg cta att cta gat gac aat cca ata acc aga att   672
His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg   720
Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

gtt aat aac tac tta gaa gct ctt ccc aag cag atg tgt gcc caa atg   768
Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc   816
Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca gtg ctg ttt ctg   864
Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Phe Leu
        275                 280                 285

cct aga aat caa att ggt ttt gtt cca gag aag aca ttt tct tca tta   912
Pro Arg Asn Gln Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu
    290                 295                 300

aaa aat tta gga gaa ctg gat ctg tct agc aat acg ata acg gag cta   960
Lys Asn Leu Gly Glu Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu
305                 310                 315                 320

tca cct cac ctt ttt aaa gac ttg aag ctt cta caa aag ctg aac ctg  1008
Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asn Leu
                325                 330                 335

tca tcc aat cct ctt atg tat ctt cac aag aac cag ttt gaa agt ctt  1056
Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln Phe Glu Ser Leu
```

```
                340                 345                 350

aaa caa ctt cag tct cta gac ctg gaa agg ata gag att cca aat ata   1104
Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile
        355                 360                 365

aac aca cga atg ttt caa ccc atg aag aat ctt tct cac att tat ttc   1152
Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Phe
    370                 375                 380

aaa aac ttt cga tac tgc tcc tat gct ccc cat gtc cga ata tgt atg   1200
Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val Arg Ile Cys Met
385                 390                 395                 400

ccc ttg acg gac ggc att tct tca ttt gag gac ctc ttg gct aac aat   1248
Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn Asn
                405                 410                 415

atc ctc aga ata ttt gtc tgg gtt ata gct ttc att acc tgc ttt gga   1296
Ile Leu Arg Ile Phe Val Trp Val Ile Ala Phe Ile Thr Cys Phe Gly
            420                 425                 430

aat ctt ttt gtc att ggc atg aga tct ttc att aaa gct gaa aat aca   1344
Asn Leu Phe Val Ile Gly Met Arg Ser Phe Ile Lys Ala Glu Asn Thr
        435                 440                 445

act cac gct atg tcc atc aaa atc ctt tgt tgt gct gat tgc ctg atg   1392
Thr His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met
    450                 455                 460

ggt gtt tac ttg ttc ttt gtt ggc att ttc gat ata aaa tac cga ggg   1440
Gly Val Tyr Leu Phe Phe Val Gly Ile Phe Asp Ile Lys Tyr Arg Gly
465                 470                 475                 480

cag tat cag aag tat gcc ttg ctg tgg atg gag agc gtg cag tgc cgc   1488
Gln Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser Val Gln Cys Arg
                485                 490                 495

ctc atg ggg ttc ctg gcc atg ctg tcc acc gaa gtc tct gtt ctg cta   1536
Leu Met Gly Phe Leu Ala Met Leu Ser Thr Glu Val Ser Val Leu Leu
            500                 505                 510

ctg acc tac ttg act ttg gag aag ttc ctg gtc att gtc ttc ccc ttc   1584
Leu Thr Tyr Leu Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe
        515                 520                 525

agt aac att cga cct gga aaa cgg cag acc tca gtc atc ctc att tgc   1632
Ser Asn Ile Arg Pro Gly Lys Arg Gln Thr Ser Val Ile Leu Ile Cys
    530                 535                 540

atc tgg atg gcg gga ttt tta ata gct gta att cca ttt tgg aat aag   1680
Ile Trp Met Ala Gly Phe Leu Ile Ala Val Ile Pro Phe Trp Asn Lys
545                 550                 555                 560

gat tat ttt gga aac ttt tat ggg aaa aat gga gta tgt ttc cca ctt   1728
Asp Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu
                565                 570                 575

tat tat gac caa aca gaa gat att gga agc aaa ggg tat tct ctt gga   1776
Tyr Tyr Asp Gln Thr Glu Asp Ile Gly Ser Lys Gly Tyr Ser Leu Gly
            580                 585                 590
```

```
att ttc cta ggt gtg aac ttg ctg gct ttt ctc atc att gtg ttt tcc   1824
Ile Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Ile Ile Val Phe Ser
        595                 600                 605

tat att act atg ttc tgt tcc att caa aaa acc gcc ttg cag acc aca   1872
Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys Thr Ala Leu Gln Thr Thr
    610                 615                 620

gaa gta agg aat tgt ttt gga aga gag gtg gct gtt gca aat cgt ttc   1920
Glu Val Arg Asn Cys Phe Gly Arg Glu Val Ala Val Ala Asn Arg Phe
625                 630                 635                 640

ttt ttt ata gtg ttc tct gat gcc atc tgc tgg att cct gta ttt gta   1968
Phe Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val
                645                 650                 655

gtt aaa atc ctt tcc ctc ttc cgg gtg gaa ata cca gac aca atg act   2016
Val Lys Ile Leu Ser Leu Phe Arg Val Glu Ile Pro Asp Thr Met Thr
            660                 665                 670

tcc tgg ata gtg att ttt ttc ctt cca gtt aac agt gct ttg aat cca   2064
Ser Trp Ile Val Ile Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro
        675                 680                 685

atc ctc tat act ctc aca acc aac ttt ttt aag gac aag ttg aaa cag   2112
Ile Leu Tyr Thr Leu Thr Thr Asn Phe Phe Lys Asp Lys Leu Lys Gln
    690                 695                 700

ctg ctg cac aaa cat cag agg aaa tca att ttc aaa att aaa aaa aaa   2160
Leu Leu His Lys His Gln Arg Lys Ser Ile Phe Lys Ile Lys Lys Lys
705                 710                 715                 720

agt tta tct aca tcc att gtg tgg ata gag gac tcc tct tcc ctg aaa   2208
Ser Leu Ser Thr Ser Ile Val Trp Ile Glu Asp Ser Ser Ser Leu Lys
                725                 730                 735

ctt ggg gtt ttg aac aaa ata aca ctt gga gac agt ata atg aaa cca   2256
Leu Gly Val Leu Asn Lys Ile Thr Leu Gly Asp Ser Ile Met Lys Pro
            740                 745                 750

gtt tcc tag                                                     2265
Val Ser
```

<210> 2
<211> 754
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
             20                  25                  30

Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
         35                  40                  45

Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
```

```
     50                  55                  60

Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
 65                  70                  75                  80

Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                 85                  90                  95

Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205

His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Phe Leu
        275                 280                 285

Pro Arg Asn Gln Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu
    290                 295                 300

Lys Asn Leu Gly Glu Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu
305                 310                 315                 320

Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asn Leu
                325                 330                 335

Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln Phe Glu Ser Leu
            340                 345                 350

Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile
        355                 360                 365

Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Phe
    370                 375                 380
```

Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val Arg Ile Cys Met
385 390 395 400

Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn Asn
405 410 415

Ile Leu Arg Ile Phe Val Trp Val Ile Ala Phe Ile Thr Cys Phe Gly
420 425 430

Asn Leu Phe Val Ile Gly Met Arg Ser Phe Ile Lys Ala Glu Asn Thr
435 440 445

Thr His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met
450 455 460

Gly Val Tyr Leu Phe Phe Val Gly Ile Phe Asp Ile Lys Tyr Arg Gly
465 470 475 480

Gln Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser Val Gln Cys Arg
485 490 495

Leu Met Gly Phe Leu Ala Met Leu Ser Thr Glu Val Ser Val Leu Leu
500 505 510

Leu Thr Tyr Leu Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe
515 520 525

Ser Asn Ile Arg Pro Gly Lys Arg Gln Thr Ser Val Ile Leu Ile Cys
530 535 540

Ile Trp Met Ala Gly Phe Leu Ile Ala Val Ile Pro Phe Trp Asn Lys
545 550 555 560

Asp Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu
565 570 575

Tyr Tyr Asp Gln Thr Glu Asp Ile Gly Ser Lys Gly Tyr Ser Leu Gly
580 585 590

Ile Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Ile Ile Val Phe Ser
595 600 605

Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys Thr Ala Leu Gln Thr Thr
610 615 620

Glu Val Arg Asn Cys Phe Gly Arg Glu Val Ala Val Ala Asn Arg Phe
625 630 635 640

Phe Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val
645 650 655

Val Lys Ile Leu Ser Leu Phe Arg Val Glu Ile Pro Asp Thr Met Thr
660 665 670

Ser Trp Ile Val Ile Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro
675 680 685

Ile Leu Tyr Thr Leu Thr Thr Asn Phe Phe Lys Asp Lys Leu Lys Gln
690 695 700

```
Leu Leu His Lys His Gln Arg Lys Ser Ile Phe Lys Ile Lys Lys Lys
705                 710                 715                 720

Ser Leu Ser Thr Ser Ile Val Trp Ile Glu Asp Ser Ser Ser Leu Lys
                725                 730                 735

Leu Gly Val Leu Asn Lys Ile Thr Leu Gly Asp Ser Ile Met Lys Pro
            740                 745                 750

Val Ser
```

<210> 3
<211> 718
<212> PRT
<213> Homo sapiens

<400> 3

```
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205

Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220
```

Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225 230 235 240

Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
245 250 255

Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
260 265 270

Glu Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
275 280 285

Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
290 295 300

Leu Met Tyr Leu His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln
305 310 315 320

Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met
325 330 335

Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg
340 345 350

Tyr Cys Ser Tyr Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp
355 360 365

Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg Ile
370 375 380

Phe Val Trp Val Ile Ala Phe Ile Thr Cys Phe Gly Asn Leu Phe Val
385 390 395 400

Ile Gly Met Arg Ser Phe Ile Lys Ala Glu Asn Thr Thr His Ala Met
405 410 415

Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu
420 425 430

Phe Phe Val Gly Ile Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys
435 440 445

Tyr Ala Leu Leu Trp Met Glu Ser Val Gln Cys Arg Leu Met Gly Phe
450 455 460

Leu Ala Met Leu Ser Thr Glu Val Ser Val Leu Leu Leu Thr Tyr Leu
465 470 475 480

Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe Ser Asn Ile Arg
485 490 495

Pro Gly Lys Arg Gln Thr Ser Val Ile Leu Ile Cys Ile Trp Met Ala
500 505 510

Gly Phe Leu Ile Ala Val Ile Pro Phe Trp Asn Lys Asp Tyr Phe Gly
515 520 525

Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu Tyr Tyr Asp Gln
530 535 540

Thr Glu Asp Ile Gly Ser Lys Gly Tyr Ser Leu Gly Ile Phe Leu Gly

```
545                 550                 555                 560

Val Asn Leu Leu Ala Phe Leu Ile Ile Val Phe Ser Tyr Ile Thr Met
                565                 570                 575

Phe Cys Ser Ile Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn
            580                 585                 590

Cys Phe Gly Arg Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val
        595                 600                 605

Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu
    610                 615                 620

Ser Leu Phe Arg Val Glu Ile Pro Asp Thr Met Thr Ser Trp Ile Val
625                 630                 635                 640

Ile Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr
                645                 650                 655

Leu Thr Thr Asn Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys
            660                 665                 670

His Gln Arg Lys Ser Ile Phe Lys Ile Lys Lys Lys Ser Leu Ser Thr
        675                 680                 685

Ser Ile Val Trp Ile Glu Asp Ser Ser Ser Leu Lys Leu Gly Val Leu
    690                 695                 700

Asn Lys Ile Thr Leu Gly Asp Ser Ile Met Lys Pro Val Ser
705                 710                 715
```

<210> 4
<211> 1149
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1149)

<400> 4

```
caa ggt agc atg atc act cct tca tgc caa aaa gga tat ttt ccc tgt   48
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctt acc aag tgc tta ccc cga gct ttt cac tgt gat ggc aag   96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
             20                  25                  30

gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
         35                  40                  45

gga tgg gcg acc ata ttt ggc aca gtg cat gga aat gct aac agc gtg   192
Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
     50                  55                  60

gcc tta aca cag gag tgc ttt cta aaa cag tat cca caa tgc tgt gac   240
Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
```

```
65                  70                  75                  80

tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac tta aag tct gtg   288
Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

ccg atg att tct aac aat gtg aca tta ctg tct ctt aag aaa aac aaa   336
Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac agt ctt cca gat aaa gtt ttc atc aaa tac aca aaa ctt aaa   384
Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

aag ata ttt ctt cag cat aat tgc att aga cac ata tcc agg aaa gca   432
Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

ttt ttt gga tta tgt aat ctg caa ata tta tat ctc aac cac aac tgc   480
Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

atc aca acc ctc aga cct gga ata ttc aaa gac tta cat cag cta act   528
Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

tgg cta att cta gat gac aat cca ata acc aga att tca cag cgc ttg   576
Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg gtt aat aac tac   624
Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205

tta gaa gct ctt ccc aag cag atg tgt gcc caa atg cct caa ctc aac   672
Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc aca aat tct acg   720
Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225                 230                 235                 240

ttt ctg tcg tgc gat tcg ctc aca gtg ctg ttt ctg cct aga aat caa   768
Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
                245                 250                 255

att ggt ttt gtt cca gag aag aca ttt tct tca tta aaa aat tta gga   816
Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
            260                 265                 270

gaa ctg gat ctg tct agc aat acg ata acg gag cta tca cct cac ctt   864
Glu Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
        275                 280                 285

ttt aaa gac ttg aag ctt cta caa aag ctg aac ctg tca tcc aat cct   912
Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
    290                 295                 300

ctt atg tat ctt cac aag aac cag ttt gaa agt ctt aaa caa ctt cag   960
Leu Met Tyr Leu His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln
305                 310                 315                 320
```

```
tct cta gac ctg gaa agg ata gag att cca aat ata aac aca cga atg   1008
Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met
                325                 330                 335

ttt caa ccc atg aag aat ctt tct cac att tat ttc aaa aac ttt cga   1056
Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg
            340                 345                 350

tac tgc tcc tat gct ccc cat gtc cga ata tgt atg ccc ttg acg gac   1104
Tyr Cys Ser Tyr Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp
        355                 360                 365

ggc att tct tca ttt gag gac ctc ttg gct aac aat atc ctc aga       1149
Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg
    370                 375                 380
```

<210> 5
<211> 383
<212> PRT
<213> Homo sapiens

<400> 5

```
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
```

```
        195                 200                 205

Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225                 230                 235                 240

Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
                245                 250                 255

Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
            260                 265                 270

Glu Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
        275                 280                 285

Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
    290                 295                 300

Leu Met Tyr Leu His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln
305                 310                 315                 320

Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met
                325                 330                 335

Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg
            340                 345                 350

Tyr Cys Ser Tyr Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp
        355                 360                 365

Gly Ile Ser Ser Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg
    370                 375                 380
```

<210> 6
<211> 2193
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(2190)

<220>
<221> sig_peptide
<222> (1)..(108)

<400> 6

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att   48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat   96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
             20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga   144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
         35                  40                  45
```

```
tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac   192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
     50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt   240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
 65                  70                  75                  80

ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat   288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                 85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca   336
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac   384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt   432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac   480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata   528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta tat ctc   576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

aac cac aac tgc atc aca acc ctc aga cct gga ata ttc aaa gac tta   624
Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205

cat cag cta act tgg cta att cta gat gac aat cca ata acc aga att   672
His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg   720
Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

gtt aat aac tac tta gaa gct ctt ccc aag cag atg tgt gcc caa atg   768
Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc   816
Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca gtg ctg gat ctg   864
Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu
        275                 280                 285
```

```
tct agc aat acg ata acg gag cta tca cct cac ctt ttt aaa gac ttg   912
Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu
    290                 295                 300

aag ctt cta caa aag ctg aac ctg tca tcc aat cct ctt atg tat ctt   960
Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu
305                 310                 315                 320

cac aag aac cag ttt gaa agt ctt aaa caa ctt cag tct cta gac ctg   1008
His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu
                325                 330                 335

gaa agg ata gag att cca aat ata aac aca cga atg ttt caa ccc atg   1056
Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
            340                 345                 350

aag aat ctt tct cac att tat ttc aaa aac ttt cga tac tgc tcc tat   1104
Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr
        355                 360                 365

gct ccc cat gtc cga ata tgt atg ccc ttg acg gac ggc att tct tca   1152
Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser
    370                 375                 380

ttt gag gac ctc ttg gct aac aat atc ctc aga ata ttt gtc tgg gtt   1200
Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val
385                 390                 395                 400

ata gct ttc att acc tgc ttt gga aat ctt ttt gtc att ggc atg aga   1248
Ile Ala Phe Ile Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg
                405                 410                 415

tct ttc att aaa gct gaa aat aca act cac gct atg tcc atc aaa atc   1296
Ser Phe Ile Lys Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile
            420                 425                 430

ctt tgt tgt gct gat tgc ctg atg ggt gtt tac ttg ttc ttt gtt ggc   1344
Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly
        435                 440                 445

att ttc gat ata aaa tac cga ggg cag tat cag aag tat gcc ttg ctg   1392
Ile Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu
    450                 455                 460

tgg atg gag agc gtg cag tgc cgc ctc atg ggg ttc ctg gcc atg ctg   1440
Trp Met Glu Ser Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu
465                 470                 475                 480

tcc acc gaa gtc tct gtt ctg cta ctg acc tac ttg act ttg gag aag   1488
Ser Thr Glu Val Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys
                485                 490                 495

ttc ctg gtc att gtc ttc ccc ttc agt aac att cga cct gga aaa cgg   1536
Phe Leu Val Ile Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg
            500                 505                 510

cag acc tca gtc atc ctc att tgc atc tgg atg gcg gga ttt tta ata   1584
Gln Thr Ser Val Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile
        515                 520                 525

gct gta att cca ttt tgg aat aag gat tat ttt gga aac ttt tat ggg   1632
```

```
Ala Val Ile Pro Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly
    530                 535                 540

aaa aat gga gta tgt ttc cca ctt tat tat gac caa aca gaa gat att   1680
Lys Asn Gly Val Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile
545                 550                 555                 560

gga agc aaa ggg tat tct ctt gga att ttc cta ggt gtg aac ttg ctg   1728
Gly Ser Lys Gly Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu
                565                 570                 575

gct ttt ctc atc att gtg ttt tcc tat att act atg ttc tgt tcc att   1776
Ala Phe Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile
            580                 585                 590

caa aaa acc gcc ttg cag acc aca gaa gta agg aat tgt ttt gga aga   1824
Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg
        595                 600                 605

gag gtg gct gtt gca aat cgt ttc ttt ttt ata gtg ttc tct gat gcc   1872
Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala
    610                 615                 620

atc tgc tgg att cct gta ttt gta gtt aaa atc ctt tcc ctc ttc cgg   1920
Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg
625                 630                 635                 640

gtg gaa ata cca gac aca atg act tcc tgg ata gtg att ttt ttc ctt   1968
Val Glu Ile Pro Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu
                645                 650                 655

cca gtt aac agt gct ttg aat cca atc ctc tat act ctc aca acc aac   2016
Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn
            660                 665                 670

ttt ttt aag gac aag ttg aaa cag ctg ctg cac aaa cat cag agg aaa   2064
Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys
        675                 680                 685

tca att ttc aaa att aaa aaa aaa agt tta tct aca tcc att gtg tgg   2112
Ser Ile Phe Lys Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp
    690                 695                 700

ata gag gac tcc tct tcc ctg aaa ctt ggg gtt ttg aac aaa ata aca   2160
Ile Glu Asp Ser Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr
705                 710                 715                 720

ctt gga gac agt ata atg aaa cca gtt tcc tag                      2193
Leu Gly Asp Ser Ile Met Lys Pro Val Ser
                725                 730


<210> 7
<211> 730
<212> PRT
<213> Homo sapiens

<400> 7
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15
```

```
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
            20                  25                  30

Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
        35                  40                  45

Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
    50                  55                  60

Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
65                  70                  75                  80

Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                85                  90                  95

Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205

His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu
        275                 280                 285

Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu
    290                 295                 300

Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu
305                 310                 315                 320

His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu
                325                 330                 335

Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
```

```
            340                 345                 350

Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr
        355                 360                 365

Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser
    370                 375                 380

Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val
385                 390                 395                 400

Ile Ala Phe Ile Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg
                405                 410                 415

Ser Phe Ile Lys Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile
            420                 425                 430

Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly
        435                 440                 445

Ile Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu
    450                 455                 460

Trp Met Glu Ser Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu
465                 470                 475                 480

Ser Thr Glu Val Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys
                485                 490                 495

Phe Leu Val Ile Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg
            500                 505                 510

Gln Thr Ser Val Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile
        515                 520                 525

Ala Val Ile Pro Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly
    530                 535                 540

Lys Asn Gly Val Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile
545                 550                 555                 560

Gly Ser Lys Gly Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu
                565                 570                 575

Ala Phe Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile
            580                 585                 590

Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg
        595                 600                 605

Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala
    610                 615                 620

Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg
625                 630                 635                 640

Val Glu Ile Pro Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu
                645                 650                 655

Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn
            660                 665                 670
```

```
Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys
        675                 680                 685

Ser Ile Phe Lys Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp
    690                 695                 700

Ile Glu Asp Ser Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr
705                 710                 715                 720

Leu Gly Asp Ser Ile Met Lys Pro Val Ser
                725                 730
```

<210> 8
<211> 694
<212> PRT
<213> *Homo sapiens*

<400> 8

```
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
             20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
         35                  40                  45

Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
     50                  55                  60

Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
 65                  70                  75                  80

Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205

Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220
```

Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225 230 235 240

Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr
245 250 255

Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln
260 265 270

Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln
275 280 285

Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu
290 295 300

Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
305 310 315 320

His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
325 330 335

Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
340 345 350

Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val Ile Ala Phe Ile
355 360 365

Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg Ser Phe Ile Lys
370 375 380

Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala
385 390 395 400

Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly Ile Phe Asp Ile
405 410 415

Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser
420 425 430

Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu Ser Thr Glu Val
435 440 445

Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys Phe Leu Val Ile
450 455 460

Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg Gln Thr Ser Val
465 470 475 480

Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile Ala Val Ile Pro
485 490 495

Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val
500 505 510

Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile Gly Ser Lys Gly
515 520 525

Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Ile
530 535 540

```
Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys Thr Ala
545                 550                 555                 560

Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg Glu Val Ala Val
                565                 570                 575

Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile
            580                 585                 590

Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg Val Glu Ile Pro
        595                 600                 605

Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu Pro Val Asn Ser
    610                 615                 620

Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn Phe Phe Lys Asp
625                 630                 635                 640

Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys Ser Ile Phe Lys
                645                 650                 655

Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp Ile Glu Asp Ser
            660                 665                 670

Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr Leu Gly Asp Ser
        675                 680                 685

Ile Met Lys Pro Val Ser
    690
```

<210> 9
<211> 1077
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1077)

<400> 9

```
caa ggt agc atg atc act cct tca tgc caa aaa gga tat ttt ccc tgt   48
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctt acc aag tgc tta ccc cga gct ttt cac tgt gat ggc aag   96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
             20                  25                  30

gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
         35                  40                  45

gga tgg gcg acc ata ttt ggc aca gtg cat gga aat gct aac agc gtg   192
Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
     50                  55                  60

gcc tta aca cag gag tgc ttt cta aaa cag tat cca caa tgc tgt gac   240
Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
 65                  70                  75                  80
```

```
tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac tta aag tct gtg   288
Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

ccg atg att tct aac aat gtg aca tta ctg tct ctt aag aaa aac aaa   336
Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac agt ctt cca gat aaa gtt ttc atc aaa tac aca aaa ctt aaa   384
Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

aag ata ttt ctt cag cat aat tgc att aga cac ata tcc agg aaa gca   432
Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

ttt ttt gga tta tgt aat ctg caa ata tta tat ctc aac cac aac tgc   480
Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

atc aca acc ctc aga cct gga ata ttc aaa gac tta cat cag cta act   528
Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

tgg cta att cta gat gac aat cca ata acc aga att tca cag cgc ttg   576
Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg gtt aat aac tac   624
Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205

tta gaa gct ctt ccc aag cag atg tgt gcc caa atg cct caa ctc aac   672
Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc aca aat tct acg   720
Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225                 230                 235                 240

ttt ctg tcg tgc gat tcg ctc aca gtg ctg gat ctg tct agc aat acg   768
Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr
                245                 250                 255

ata acg gag cta tca cct cac ctt ttt aaa gac ttg aag ctt cta caa   816
Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln
            260                 265                 270

aag ctg aac ctg tca tcc aat cct ctt atg tat ctt cac aag aac cag   864
Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln
        275                 280                 285

ttt gaa agt ctt aaa caa ctt cag tct cta gac ctg gaa agg ata gag   912
Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu
    290                 295                 300

att cca aat ata aac aca cga atg ttt caa ccc atg aag aat ctt tct   960
Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
305                 310                 315                 320

cac att tat ttc aaa aac ttt cga tac tgc tcc tat gct ccc cat gtc   1008
```

```
His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
                325                 330                 335

cga ata tgt atg ccc ttg acg gac ggc att tct tca ttt gag gac ctc   1056
Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
            340                 345                 350

ttg gct aac aat atc ctc aga                                       1077
Leu Ala Asn Asn Ile Leu Arg
        355
```

<210> 10
<211> 359
<212> PRT
<213> Homo sapiens

```
<400> 10
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205

Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
```

```
225                 230                 235                 240

Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr
                245                 250                 255

Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln
            260                 265                 270

Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln
        275                 280                 285

Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu
    290                 295                 300

Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
305                 310                 315                 320

His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
                325                 330                 335

Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
            340                 345                 350

Leu Ala Asn Asn Ile Leu Arg
        355
```

<210> 11
<211> 2049
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(2046)

<220>
<221> sig_peptide
<222> (1)..(108)

<400> 11

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att   48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat   96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
            20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga   144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
        35                  40                  45

tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac   192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
    50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt   240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
 65                  70                  75                  80
```

```
ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat   288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                 85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca   336
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac   384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt   432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac   480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata   528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta att cta   576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Ile Leu
            180                 185                 190

gat gac aat cca ata acc aga att tca cag cgc ttg ttt acg gga tta   624
Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu Phe Thr Gly Leu
        195                 200                 205

aat tcc ttg ttt ttc ctg tct atg gtt aat aac tac tta gaa gct ctt   672
Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr Leu Glu Ala Leu
    210                 215                 220

ccc aag cag atg tgt gcc caa atg cct caa ctc aac tgg gtg gat ttg   720
Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn Trp Val Asp Leu
225                 230                 235                 240

gaa ggc aat aga ata aag tat ctc aca aat tct acg ttt ctg tcg tgc   768
Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr Phe Leu Ser Cys
                245                 250                 255

gat tcg ctc aca gtg ctg gat ctg tct agc aat acg ata acg gag cta   816
Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu
            260                 265                 270

tca cct cac ctt ttt aaa gac ttg aag ctt cta caa aag cta gac ctg   864
Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asp Leu
        275                 280                 285

gaa agg ata gag att cca aat ata aac aca cga atg ttt caa ccc atg   912
Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
    290                 295                 300

aag aat ctt tct cac att tat ttc aaa aac ttt cga tac tgc tcc tat   960
Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr
305                 310                 315                 320

gct ccc cat gtc cga ata tgt atg ccc ttg acg gac ggc att tct tca   1008
```

```
Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser
                325                 330                 335

ttt gag gac ctc ttg gct aac aat atc ctc aga ata ttt gtc tgg gtt   1056
Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val
            340                 345                 350

ata gct ttc att acc tgc ttt gga aat ctt ttt gtc att ggc atg aga   1104
Ile Ala Phe Ile Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg
        355                 360                 365

tct ttc att aaa gct gaa aat aca act cac gct atg tcc atc aaa atc   1152
Ser Phe Ile Lys Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile
    370                 375                 380

ctt tgt tgt gct gat tgc ctg atg ggt gtt tac ttg ttc ttt gtt ggc   1200
Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly
385                 390                 395                 400

att ttc gat ata aaa tac cga ggg cag tat cag aag tat gcc ttg ctg   1248
Ile Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu
                405                 410                 415

tgg atg gag agc gtg cag tgc cgc ctc atg ggg ttc ctg gcc atg ctg   1296
Trp Met Glu Ser Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu
            420                 425                 430

tcc acc gaa gtc tct gtt ctg cta ctg acc tac ttg act ttg gag aag   1344
Ser Thr Glu Val Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys
        435                 440                 445

ttc ctg gtc att gtc ttc ccc ttc agt aac att cga cct gga aaa cgg   1392
Phe Leu Val Ile Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg
    450                 455                 460

cag acc tca gtc atc ctc att tgc atc tgg atg gcg gga ttt tta ata   1440
Gln Thr Ser Val Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile
465                 470                 475                 480

gct gta att cca ttt tgg aat aag gat tat ttt gga aac ttt tat ggg   1488
Ala Val Ile Pro Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly
                485                 490                 495

aaa aat gga gta tgt ttc cca ctt tat tat gac caa aca gaa gat att   1536
Lys Asn Gly Val Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile
            500                 505                 510

gga agc aaa ggg tat tct ctt gga att ttc cta ggt gtg aac ttg ctg   1584
Gly Ser Lys Gly Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu
        515                 520                 525

gct ttt ctc atc att gtg ttt tcc tat att act atg ttc tgt tcc att   1632
Ala Phe Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile
    530                 535                 540

caa aaa acc gcc ttg cag acc aca gaa gta agg aat tgt ttt gga aga   1680
Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg
545                 550                 555                 560

gag gtg gct gtt gca aat cgt ttc ttt ttt ata gtg ttc tct gat gcc   1728
Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala
```

```
                565                 570                 575

atc tgc tgg att cct gta ttt gta gtt aaa atc ctt tcc ctc ttc cgg   1776
Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg
            580                 585                 590

gtg gaa ata cca gac aca atg act tcc tgg ata gtg att ttt ttc ctt   1824
Val Glu Ile Pro Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu
        595                 600                 605

cca gtt aac agt gct ttg aat cca atc ctc tat act ctc aca acc aac   1872
Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn
    610                 615                 620

ttt ttt aag gac aag ttg aaa cag ctg ctg cac aaa cat cag agg aaa   1920
Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys
625                 630                 635                 640

tca att ttc aaa att aaa aaa aaa agt tta tct aca tcc att gtg tgg   1968
Ser Ile Phe Lys Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp
                645                 650                 655

ata gag gac tcc tct tcc ctg aaa ctt ggg gtt ttg aac aaa ata aca   2016
Ile Glu Asp Ser Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr
            660                 665                 670

ctt gga gac agt ata atg aaa cca gtt tcc tag                      2049
Leu Gly Asp Ser Ile Met Lys Pro Val Ser
        675                 680
```

<210> 12
<211> 682
<212> PRT
<213> Homo sapiens

<400> 12

```
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
             20                  25                  30

Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
         35                  40                  45

Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
     50                  55                  60

Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
 65                  70                  75                  80

Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                 85                  90                  95

Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125
```

```
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Ile Leu
            180                 185                 190

Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu Phe Thr Gly Leu
        195                 200                 205

Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr Leu Glu Ala Leu
    210                 215                 220

Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn Trp Val Asp Leu
225                 230                 235                 240

Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr Phe Leu Ser Cys
                245                 250                 255

Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu
            260                 265                 270

Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asp Leu
        275                 280                 285

Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
    290                 295                 300

Lys Asn Leu Ser His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr
305                 310                 315                 320

Ala Pro His Val Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser
                325                 330                 335

Phe Glu Asp Leu Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val
            340                 345                 350

Ile Ala Phe Ile Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg
        355                 360                 365

Ser Phe Ile Lys Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile
    370                 375                 380

Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly
385                 390                 395                 400

Ile Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu
                405                 410                 415

Trp Met Glu Ser Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu
            420                 425                 430

Ser Thr Glu Val Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys
        435                 440                 445

Phe Leu Val Ile Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg
```

```
    450                 455                 460

Gln Thr Ser Val Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile
465                 470                 475                 480

Ala Val Ile Pro Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly
                485                 490                 495

Lys Asn Gly Val Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile
            500                 505                 510

Gly Ser Lys Gly Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu
        515                 520                 525

Ala Phe Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile
    530                 535                 540

Gln Lys Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg
545                 550                 555                 560

Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala
                565                 570                 575

Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg
            580                 585                 590

Val Glu Ile Pro Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu
        595                 600                 605

Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn
    610                 615                 620

Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys
625                 630                 635                 640

Ser Ile Phe Lys Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp
                645                 650                 655

Ile Glu Asp Ser Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr
            660                 665                 670

Leu Gly Asp Ser Ile Met Lys Pro Val Ser
        675                 680


<210> 13
<211> 646
<212> PRT
<213> Homo sapiens

<400> 13
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
             20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
         35                  40                  45

Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
```

```
     50                  55                  60

Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
 65                  70                  75                  80

Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                 85                  90                  95

Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Ile Leu Asp Asp Asn Pro
145                 150                 155                 160

Ile Thr Arg Ile Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe
                165                 170                 175

Phe Leu Ser Met Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met
            180                 185                 190

Cys Ala Gln Met Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg
        195                 200                 205

Ile Lys Tyr Leu Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr
    210                 215                 220

Val Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
225                 230                 235                 240

Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asp Leu Glu Arg Ile Glu
                245                 250                 255

Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
            260                 265                 270

His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
        275                 280                 285

Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
    290                 295                 300

Leu Ala Asn Asn Ile Leu Arg Ile Phe Val Trp Val Ile Ala Phe Ile
305                 310                 315                 320

Thr Cys Phe Gly Asn Leu Phe Val Ile Gly Met Arg Ser Phe Ile Lys
                325                 330                 335

Ala Glu Asn Thr Thr His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala
            340                 345                 350

Asp Cys Leu Met Gly Val Tyr Leu Phe Phe Val Gly Ile Phe Asp Ile
        355                 360                 365

Lys Tyr Arg Gly Gln Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser
    370                 375                 380
```

```
Val Gln Cys Arg Leu Met Gly Phe Leu Ala Met Leu Ser Thr Glu Val
385                 390                 395                 400

Ser Val Leu Leu Leu Thr Tyr Leu Thr Leu Glu Lys Phe Leu Val Ile
                405                 410                 415

Val Phe Pro Phe Ser Asn Ile Arg Pro Gly Lys Arg Gln Thr Ser Val
            420                 425                 430

Ile Leu Ile Cys Ile Trp Met Ala Gly Phe Leu Ile Ala Val Ile Pro
        435                 440                 445

Phe Trp Asn Lys Asp Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val
    450                 455                 460

Cys Phe Pro Leu Tyr Tyr Asp Gln Thr Glu Asp Ile Gly Ser Lys Gly
465                 470                 475                 480

Tyr Ser Leu Gly Ile Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Ile
                485                 490                 495

Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys Thr Ala
            500                 505                 510

Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg Glu Val Ala Val
        515                 520                 525

Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile
    530                 535                 540

Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg Val Glu Ile Pro
545                 550                 555                 560

Asp Thr Met Thr Ser Trp Ile Val Ile Phe Phe Leu Pro Val Asn Ser
                565                 570                 575

Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Asn Phe Phe Lys Asp
            580                 585                 590

Lys Leu Lys Gln Leu Leu His Lys His Gln Arg Lys Ser Ile Phe Lys
        595                 600                 605

Ile Lys Lys Lys Ser Leu Ser Thr Ser Ile Val Trp Ile Glu Asp Ser
    610                 615                 620

Ser Ser Leu Lys Leu Gly Val Leu Asn Lys Ile Thr Leu Gly Asp Ser
625                 630                 635                 640

Ile Met Lys Pro Val Ser
                645
```

<210> 14
<211> 933
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(933)

<400> 14

```
caa ggt agc atg atc act cct tca tgc caa aaa gga tat ttt ccc tgt   48
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctt acc aag tgc tta ccc cga gct ttt cac tgt gat ggc aag   96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

gga tgg gcg acc ata ttt ggc aca gtg cat gga aat gct aac agc gtg   192
Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

gcc tta aca cag gag tgc ttt cta aaa cag tat cca caa tgc tgt gac   240
Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac tta aag tct gtg   288
Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

ccg atg att tct aac aat gtg aca tta ctg tct ctt aag aaa aac aaa   336
Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac agt ctt cca gat aaa gtt ttc atc aaa tac aca aaa ctt aaa   384
Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

aag ata ttt ctt cag cat aat tgc att aga cac ata tcc agg aaa gca   432
Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

ttt ttt gga tta tgt aat ctg caa ata tta att cta gat gac aat cca   480
Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Ile Leu Asp Asp Asn Pro
145                 150                 155                 160

ata acc aga att tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt   528
Ile Thr Arg Ile Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe
                165                 170                 175

ttc ctg tct atg gtt aat aac tac tta gaa gct ctt ccc aag cag atg   576
Phe Leu Ser Met Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met
            180                 185                 190

tgt gcc caa atg cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga   624
Cys Ala Gln Met Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg
        195                 200                 205

ata aag tat ctc aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca   672
Ile Lys Tyr Leu Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr
    210                 215                 220

gtg ctg gat ctg tct agc aat acg ata acg gag cta tca cct cac ctt   720
Val Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
225                 230                 235                 240
```

```
ttt aaa gac ttg aag ctt cta caa aag cta gac ctg gaa agg ata gag   768
Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asp Leu Glu Arg Ile Glu
                245                 250                 255

att cca aat ata aac aca cga atg ttt caa ccc atg aag aat ctt tct   816
Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
            260                 265                 270

cac att tat ttc aaa aac ttt cga tac tgc tcc tat gct ccc cat gtc   864
His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
        275                 280                 285

cga ata tgt atg ccc ttg acg gac ggc att tct tca ttt gag gac ctc   912
Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
    290                 295                 300

ttg gct aac aat atc ctc aga                                      933
Leu Ala Asn Asn Ile Leu Arg
305                 310
```

<210> 15
<211> 311
<212> PRT
<213> Homo sapiens

<400> 15

```
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140

Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Ile Leu Asp Asp Asn Pro
145                 150                 155                 160

Ile Thr Arg Ile Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe
                165                 170                 175
```

```
Phe Leu Ser Met Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met
            180                 185                 190

Cys Ala Gln Met Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg
        195                 200                 205

Ile Lys Tyr Leu Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr
    210                 215                 220

Val Leu Asp Leu Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu
225                 230                 235                 240

Phe Lys Asp Leu Lys Leu Leu Gln Lys Leu Asp Leu Glu Arg Ile Glu
                245                 250                 255

Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
            260                 265                 270

His Ile Tyr Phe Lys Asn Phe Arg Tyr Cys Ser Tyr Ala Pro His Val
        275                 280                 285

Arg Ile Cys Met Pro Leu Thr Asp Gly Ile Ser Ser Phe Glu Asp Leu
    290                 295                 300

Leu Ala Asn Asn Ile Leu Arg
305                 310
```

<210> 16
<211> 1101
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1098)

<220>
<221> sig_peptide
<222> (1)..(108)

<400> 16

```
atg att gtt ttt ctg gtt ttt aaa cat ctc ttc agc ctc aga ttg att   48
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

aca atg ttc ttt cta ctt cat ttc atc gtt ctg atc aat gtc aaa gat   96
Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
             20                  25                  30

ttt gca ctg act caa ggt agc atg atc act cct tca tgc caa aaa gga   144
Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
         35                  40                  45

tat ttt ccc tgt ggg aat ctt acc aag tgc tta ccc cga gct ttt cac   192
Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
     50                  55                  60

tgt gat ggc aag gat gac tgt ggg aac ggg gcg gac gaa gag aac tgt   240
Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
 65                  70                  75                  80
```

```
ggt gac act agt gga tgg gcg acc ata ttt ggc aca gtg cat gga aat   288
Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                 85                  90                  95

gct aac agc gtg gcc tta aca cag gag tgc ttt cta aaa cag tat cca   336
Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

caa tgc tgt gac tgc aaa gaa act gaa ttg gaa tgt gta aat ggt gac   384
Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

tta aag tct gtg ccg atg att tct aac aat gtg aca tta ctg tct ctt   432
Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

aag aaa aac aaa atc cac agt ctt cca gat aaa gtt ttc atc aaa tac   480
Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

aca aaa ctt aaa aag ata ttt ctt cag cat aat tgc att aga cac ata   528
Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

tcc agg aaa gca ttt ttt gga tta tgt aat ctg caa ata tta tat ctc   576
Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

aac cac aac tgc atc aca acc ctc aga cct gga ata ttc aaa gac tta   624
Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205

cat cag cta act tgg cta att cta gat gac aat cca ata acc aga att   672
His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220

tca cag cgc ttg ttt acg gga tta aat tcc ttg ttt ttc ctg tct atg   720
Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

gtt aat aac tac tta gaa gct ctt ccc aag cag atg tgt gcc caa atg   768
Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

cct caa ctc aac tgg gtg gat ttg gaa ggc aat aga ata aag tat ctc   816
Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

aca aat tct acg ttt ctg tcg tgc gat tcg ctc aca gtg ctg gat ctg   864
Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu
        275                 280                 285

tct agc aat acg ata acg gag cta tca cct cac ctt ttt aaa gac ttg   912
Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu
    290                 295                 300

aag ctt cta caa aag ctg aac ctg tca tcc aat cct ctt atg tat ctt   960
Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu
305                 310                 315                 320
```

```
cac aag aac cag ttt gaa agt ctt aaa caa ctt cag tct cta gac ctg   1008
His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu
                325                 330                 335

gaa agg ata gag att cca aat ata aac aca cga atg ttt caa ccc atg   1056
Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
            340                 345                 350

aag aat ctt tct cac ata gtt caa tat tat gat gtg ccg aca tga       1101
Lys Asn Leu Ser His Ile Val Gln Tyr Tyr Asp Val Pro Thr
        355                 360                 365
```

<210> 17
<211> 366
<212> PRT
<213> Homo sapiens

<400> 17

```
Met Ile Val Phe Leu Val Phe Lys His Leu Phe Ser Leu Arg Leu Ile
  1               5                  10                  15

Thr Met Phe Phe Leu Leu His Phe Ile Val Leu Ile Asn Val Lys Asp
            20                  25                  30

Phe Ala Leu Thr Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly
        35                  40                  45

Tyr Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His
    50                  55                  60

Cys Asp Gly Lys Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys
65                  70                  75                  80

Gly Asp Thr Ser Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn
                85                  90                  95

Ala Asn Ser Val Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro
            100                 105                 110

Gln Cys Cys Asp Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp
        115                 120                 125

Leu Lys Ser Val Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu
    130                 135                 140

Lys Lys Asn Lys Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr
145                 150                 155                 160

Thr Lys Leu Lys Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile
                165                 170                 175

Ser Arg Lys Ala Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu
            180                 185                 190

Asn His Asn Cys Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu
        195                 200                 205

His Gln Leu Thr Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile
    210                 215                 220
```

```
Ser Gln Arg Leu Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met
225                 230                 235                 240

Val Asn Asn Tyr Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met
                245                 250                 255

Pro Gln Leu Asn Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu
            260                 265                 270

Thr Asn Ser Thr Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu
        275                 280                 285

Ser Ser Asn Thr Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu
    290                 295                 300

Lys Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu
305                 310                 315                 320

His Lys Asn Gln Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu
                325                 330                 335

Glu Arg Ile Glu Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met
            340                 345                 350

Lys Asn Leu Ser His Ile Val Gln Tyr Tyr Asp Val Pro Thr
        355                 360                 365
```

<210> 18
<211> 330
<212> PRT
<213> Homo sapiens

<400> 18

```
Gln Gly Ser Met Ile Thr Pro Ser Cys Gln Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Lys
            20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Glu Asn Cys Gly Asp Thr Ser
        35                  40                  45

Gly Trp Ala Thr Ile Phe Gly Thr Val His Gly Asn Ala Asn Ser Val
    50                  55                  60

Ala Leu Thr Gln Glu Cys Phe Leu Lys Gln Tyr Pro Gln Cys Cys Asp
65                  70                  75                  80

Cys Lys Glu Thr Glu Leu Glu Cys Val Asn Gly Asp Leu Lys Ser Val
                85                  90                  95

Pro Met Ile Ser Asn Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Ser Leu Pro Asp Lys Val Phe Ile Lys Tyr Thr Lys Leu Lys
        115                 120                 125

Lys Ile Phe Leu Gln His Asn Cys Ile Arg His Ile Ser Arg Lys Ala
    130                 135                 140
```

```
Phe Phe Gly Leu Cys Asn Leu Gln Ile Leu Tyr Leu Asn His Asn Cys
145                 150                 155                 160

Ile Thr Thr Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Thr
                165                 170                 175

Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Arg Leu
            180                 185                 190

Phe Thr Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Asn Asn Tyr
        195                 200                 205

Leu Glu Ala Leu Pro Lys Gln Met Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

Trp Val Asp Leu Glu Gly Asn Arg Ile Lys Tyr Leu Thr Asn Ser Thr
225                 230                 235                 240

Phe Leu Ser Cys Asp Ser Leu Thr Val Leu Asp Leu Ser Ser Asn Thr
                245                 250                 255

Ile Thr Glu Leu Ser Pro His Leu Phe Lys Asp Leu Lys Leu Leu Gln
            260                 265                 270

Lys Leu Asn Leu Ser Ser Asn Pro Leu Met Tyr Leu His Lys Asn Gln
        275                 280                 285

Phe Glu Ser Leu Lys Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu
    290                 295                 300

Ile Pro Asn Ile Asn Thr Arg Met Phe Gln Pro Met Lys Asn Leu Ser
305                 310                 315                 320

His Ile Val Gln Tyr Tyr Asp Val Pro Thr
                325                 330
```

<210> 19
<211> 2214
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (1)..(2211)

<220>
<221> sig_peptide
<222> (1)..(57)

<400> 19

```
atg tgg ctc cta ctt cat gtc atc ctt ctg aca gag gtc aaa gat ttt   48
Met Trp Leu Leu Leu His Val Ile Leu Leu Thr Glu Val Lys Asp Phe
  1               5                  10                  15

gca ctg gct gac agc agt atg gtg gct cct ctg tgc ccc aaa ggg tat   96
Ala Leu Ala Asp Ser Ser Met Val Ala Pro Leu Cys Pro Lys Gly Tyr
             20                  25                  30

ttt ccc tgt ggg aat ctc acc aaa tgc ttg ccc cga gcc ttt cac tgc  144
Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys
```

```
          35                  40                  45

gat ggt gtg gat gat tgc ggg aat ggt gcc gac gag gac aac tgt ggt   192
Asp Gly Val Asp Asp Cys Gly Asn Gly Ala Asp Glu Asp Asn Cys Gly
     50                  55                  60

gac act agt gga tgg aca acc ata ttt ggc aca gtc cat ggg aat gtc   240
Asp Thr Ser Gly Trp Thr Thr Ile Phe Gly Thr Val His Gly Asn Val
 65                  70                  75                  80

aat aaa gtg aca ttg aca cag gag tgc ttt ctc agc cag tat cca cag   288
Asn Lys Val Thr Leu Thr Gln Glu Cys Phe Leu Ser Gln Tyr Pro Gln
                85                  90                  95

cac tgt tac tgc aga gaa aat gaa ctg gaa tgt gta aag gct gac tta   336
His Cys Tyr Cys Arg Glu Asn Glu Leu Glu Cys Val Lys Ala Asp Leu
            100                 105                 110

aaa gct gtg cca aag gtt tcc agc aac gta aca tta cta tct ctt aag   384
Lys Ala Val Pro Lys Val Ser Ser Asn Val Thr Leu Leu Ser Leu Lys
        115                 120                 125

aaa aac aaa atc cac aga ctt cca gtc aag gtc ttc agc aga tac aca   432
Lys Asn Lys Ile His Arg Leu Pro Val Lys Val Phe Ser Arg Tyr Thr
    130                 135                 140

gaa ctc aga aag ata tac ctt cag cac aac tgc atc aca cac atc tcc   480
Glu Leu Arg Lys Ile Tyr Leu Gln His Asn Cys Ile Thr His Ile Ser
145                 150                 155                 160

agg aga gca ttc ctt gga tta cat aat cta caa ata ctg tat ctc agc   528
Arg Arg Ala Phe Leu Gly Leu His Asn Leu Gln Ile Leu Tyr Leu Ser
                165                 170                 175

cat aac tgc att acc tct ctc agg cct ggg ata ttc aaa gac ttg cat   576
His Asn Cys Ile Thr Ser Leu Arg Pro Gly Ile Phe Lys Asp Leu His
            180                 185                 190

cag ctt gct tgg cta att tta gat gac aac ccg atc acc aga atc tca   624
Gln Leu Ala Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser
        195                 200                 205

cag aag tcc ttt atg ggg tta aac tcc ttg ttt ttc ttg tcc atg gtg   672
Gln Lys Ser Phe Met Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val
    210                 215                 220

ggt aac cgg ctc gag gcc ctt cct gaa aca ttg tgt gct cag atg cct   720
Gly Asn Arg Leu Glu Ala Leu Pro Glu Thr Leu Cys Ala Gln Met Pro
225                 230                 235                 240

caa ctc aac tgg gtg gat ctg gca aac aat gga ata aag tac ata acg   768
Gln Leu Asn Trp Val Asp Leu Ala Asn Asn Gly Ile Lys Tyr Ile Thr
                245                 250                 255

aac tcc acc ttc cta acg tgc gac tcg ctc acg gtt ctg ttt ctg cct   816
Asn Ser Thr Phe Leu Thr Cys Asp Ser Leu Thr Val Leu Phe Leu Pro
            260                 265                 270

aga aat caa att ggt ttt gtt cca gag aag aca ttt tct tca tta aaa   864
Arg Asn Gln Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys
        275                 280                 285
```

```
aat tta gga gaa ctg gac ctg tct agc aat atg ata aca aaa ctc cca   912
Asn Leu Gly Glu Leu Asp Leu Ser Ser Asn Met Ile Thr Lys Leu Pro
    290                 295                 300

gtc cac ctt ttc agc gac ctt cat ctt ctc cag aag ctg aac ctg tca   960
Val His Leu Phe Ser Asp Leu His Leu Leu Gln Lys Leu Asn Leu Ser
305                 310                 315                 320

tcc aac cct ctt ctg tat gtc cac aag aac cag ttt gga agt ctc aaa  1008
Ser Asn Pro Leu Leu Tyr Val His Lys Asn Gln Phe Gly Ser Leu Lys
                325                 330                 335

caa ctt cag tct cta gac ctg gaa agg ata gag att cca aac ata agc  1056
Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Ser
            340                 345                 350

aca gga atg ttc cag cca atg aag aac ctt tct cac att tat ttg aaa  1104
Thr Gly Met Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Leu Lys
        355                 360                 365

acc ttt cga tac tgc tcc tat gtc ccc cat gtc cga atc tgt atg ccg  1152
Thr Phe Arg Tyr Cys Ser Tyr Val Pro His Val Arg Ile Cys Met Pro
    370                 375                 380

tcg act gat ggt att tct tcg tct gag gac ctc ttg gct aac ggt atc  1200
Ser Thr Asp Gly Ile Ser Ser Ser Glu Asp Leu Leu Ala Asn Gly Ile
385                 390                 395                 400

ctc aga gtg tct gtc tgg gtt ata gct ttc att acc tgc gtt ggg aat  1248
Leu Arg Val Ser Val Trp Val Ile Ala Phe Ile Thr Cys Val Gly Asn
                405                 410                 415

ttc ctt gtc ata gcc gtg aga tct ctc att aag gct gag aat aca act  1296
Phe Leu Val Ile Ala Val Arg Ser Leu Ile Lys Ala Glu Asn Thr Thr
            420                 425                 430

cac gct atg tcc atc aaa atc ctt tgt tgt gcc gat tgc ctg atg ggg  1344
His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met Gly
        435                 440                 445

gtg tac ctg ttc tcc gtg ggc gtc ttt gac atc aag tac cga ggg cag  1392
Val Tyr Leu Phe Ser Val Gly Val Phe Asp Ile Lys Tyr Arg Gly Gln
    450                 455                 460

tat cag aag tat gcg ctg ctg tgg atg gag agt gtg ccc tgc cgc ctg  1440
Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser Val Pro Cys Arg Leu
465                 470                 475                 480

ctg ggc ttc ctg gcc acg ctg tcc aca gag gtc tcg gtg ctg ctg ctg  1488
Leu Gly Phe Leu Ala Thr Leu Ser Thr Glu Val Ser Val Leu Leu Leu
                485                 490                 495

aca ttc ctg acg ctg gag aag ttc ctt gtc ata gta ttc cct ttc agc  1536
Thr Phe Leu Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe Ser
            500                 505                 510

aac ctg cgc ctg ggc aag cgc cag act gct gtg gcc ctc gcc agc atc  1584
Asn Leu Arg Leu Gly Lys Arg Gln Thr Ala Val Ala Leu Ala Ser Ile
        515                 520                 525
```

```
tgg gtg gtg gga ttt ctc ata gcg gcc gtt ccg ttc acc aga gag gat   1632
Trp Val Val Gly Phe Leu Ile Ala Ala Val Pro Phe Thr Arg Glu Asp
    530                 535                 540

tat ttc ggc aac ttt tat ggg aaa aat gga gtc tgc ttc cca ctt cat   1680
Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu His
545                 550                 555                 560

tat gac caa gca gaa gat ttt gga agt aga ggg tac tcc ctt ggg att   1728
Tyr Asp Gln Ala Glu Asp Phe Gly Ser Arg Gly Tyr Ser Leu Gly Ile
                565                 570                 575

ttc cta ggt gtg aac ttg ctg gct ttc ctc gtc atc gtg att tcc tat   1776
Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Val Ile Val Ile Ser Tyr
            580                 585                 590

gtc acc atg ttc tgc tcc att cat aaa aca gcc ctt cag act gca gaa   1824
Val Thr Met Phe Cys Ser Ile His Lys Thr Ala Leu Gln Thr Ala Glu
        595                 600                 605

gtg agg agc cac atc ggg aag gag gtg gct gtt gca aac cgg ttc ttt   1872
Val Arg Ser His Ile Gly Lys Glu Val Ala Val Ala Asn Arg Phe Phe
    610                 615                 620

ttt atc gtg ttc tct gat gcc atc tgc tgg atc cct gtg ttt gtc gtt   1920
Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val Val
625                 630                 635                 640

aag atc ctg tct ctc ctt caa gtg gag ata cca ggc aca atc act tcc   1968
Lys Ile Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr Ile Thr Ser
                645                 650                 655

tgg atc gtg gtt ttt ttc ctt ccg gtg aac agc gcc tta aac ccc atc   2016
Trp Ile Val Val Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro Ile
            660                 665                 670

ctc tac act ctg acg acc tcc ttt ttt aag gac aag ttg aaa cag ttg   2064
Leu Tyr Thr Leu Thr Thr Ser Phe Phe Lys Asp Lys Leu Lys Gln Leu
        675                 680                 685

ctg cac aaa cat cgg agg aaa ccc atc ttc aaa gtt aag aag aaa agt   2112
Leu His Lys His Arg Arg Lys Pro Ile Phe Lys Val Lys Lys Lys Ser
    690                 695                 700

tta tcc gct tcc att gtg tgg aca gac gag tct tca ctt aaa ctt gga   2160
Leu Ser Ala Ser Ile Val Trp Thr Asp Glu Ser Ser Leu Lys Leu Gly
705                 710                 715                 720

gtg ttg agc aaa ata gcc ctt ggg gac agt ata atg aag ccg gtc tcc   2208
Val Leu Ser Lys Ile Ala Leu Gly Asp Ser Ile Met Lys Pro Val Ser
                725                 730                 735

ccg tag                                                          2214
Pro
```

<210> 20
<211> 737
<212> PRT
<213> Mus musculus

<400> 20

```
Met Trp Leu Leu Leu His Val Ile Leu Leu Thr Glu Val Lys Asp Phe
  1               5                  10                  15

Ala Leu Ala Asp Ser Ser Met Val Ala Pro Leu Cys Pro Lys Gly Tyr
            20                  25                  30

Phe Pro Cys Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys
        35                  40                  45

Asp Gly Val Asp Asp Cys Gly Asn Gly Ala Asp Glu Asp Asn Cys Gly
    50                  55                  60

Asp Thr Ser Gly Trp Thr Thr Ile Phe Gly Thr Val His Gly Asn Val
65                  70                  75                  80

Asn Lys Val Thr Leu Thr Gln Glu Cys Phe Leu Ser Gln Tyr Pro Gln
                85                  90                  95

His Cys Tyr Cys Arg Glu Asn Glu Leu Glu Cys Val Lys Ala Asp Leu
            100                 105                 110

Lys Ala Val Pro Lys Val Ser Ser Asn Val Thr Leu Leu Ser Leu Lys
        115                 120                 125

Lys Asn Lys Ile His Arg Leu Pro Val Lys Val Phe Ser Arg Tyr Thr
    130                 135                 140

Glu Leu Arg Lys Ile Tyr Leu Gln His Asn Cys Ile Thr His Ile Ser
145                 150                 155                 160

Arg Arg Ala Phe Leu Gly Leu His Asn Leu Gln Ile Leu Tyr Leu Ser
                165                 170                 175

His Asn Cys Ile Thr Ser Leu Arg Pro Gly Ile Phe Lys Asp Leu His
            180                 185                 190

Gln Leu Ala Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser
        195                 200                 205

Gln Lys Ser Phe Met Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val
    210                 215                 220

Gly Asn Arg Leu Glu Ala Leu Pro Glu Thr Leu Cys Ala Gln Met Pro
225                 230                 235                 240

Gln Leu Asn Trp Val Asp Leu Ala Asn Asn Gly Ile Lys Tyr Ile Thr
                245                 250                 255

Asn Ser Thr Phe Leu Thr Cys Asp Ser Leu Thr Val Leu Phe Leu Pro
            260                 265                 270

Arg Asn Gln Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys
        275                 280                 285

Asn Leu Gly Glu Leu Asp Leu Ser Ser Asn Met Ile Thr Lys Leu Pro
    290                 295                 300

Val His Leu Phe Ser Asp Leu His Leu Leu Gln Lys Leu Asn Leu Ser
305                 310                 315                 320
```

Ser Asn Pro Leu Leu Tyr Val His Lys Asn Gln Phe Gly Ser Leu Lys
325 330 335

Gln Leu Gln Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Ser
340 345 350

Thr Gly Met Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Leu Lys
355 360 365

Thr Phe Arg Tyr Cys Ser Tyr Val Pro His Val Arg Ile Cys Met Pro
370 375 380

Ser Thr Asp Gly Ile Ser Ser Ser Glu Asp Leu Leu Ala Asn Gly Ile
385 390 395 400

Leu Arg Val Ser Val Trp Val Ile Ala Phe Ile Thr Cys Val Gly Asn
405 410 415

Phe Leu Val Ile Ala Val Arg Ser Leu Ile Lys Ala Glu Asn Thr Thr
420 425 430

His Ala Met Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met Gly
435 440 445

Val Tyr Leu Phe Ser Val Gly Val Phe Asp Ile Lys Tyr Arg Gly Gln
450 455 460

Tyr Gln Lys Tyr Ala Leu Leu Trp Met Glu Ser Val Pro Cys Arg Leu
465 470 475 480

Leu Gly Phe Leu Ala Thr Leu Ser Thr Glu Val Ser Val Leu Leu Leu
485 490 495

Thr Phe Leu Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe Ser
500 505 510

Asn Leu Arg Leu Gly Lys Arg Gln Thr Ala Val Ala Leu Ala Ser Ile
515 520 525

Trp Val Val Gly Phe Leu Ile Ala Ala Val Pro Phe Thr Arg Glu Asp
530 535 540

Tyr Phe Gly Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu His
545 550 555 560

Tyr Asp Gln Ala Glu Asp Phe Gly Ser Arg Gly Tyr Ser Leu Gly Ile
565 570 575

Phe Leu Gly Val Asn Leu Leu Ala Phe Leu Val Ile Val Ile Ser Tyr
580 585 590

Val Thr Met Phe Cys Ser Ile His Lys Thr Ala Leu Gln Thr Ala Glu
595 600 605

Val Arg Ser His Ile Gly Lys Glu Val Ala Val Ala Asn Arg Phe Phe
610 615 620

Phe Ile Val Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val Val
625 630 635 640

Lys Ile Leu Ser Leu Leu Gln Val Glu Ile Pro Gly Thr Ile Thr Ser

```
                645                 650                 655

Trp Ile Val Val Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro Ile
            660                 665                 670

Leu Tyr Thr Leu Thr Thr Ser Phe Phe Lys Asp Lys Leu Lys Gln Leu
        675                 680                 685

Leu His Lys His Arg Arg Lys Pro Ile Phe Lys Val Lys Lys Lys Ser
    690                 695                 700

Leu Ser Ala Ser Ile Val Trp Thr Asp Glu Ser Ser Leu Lys Leu Gly
705                 710                 715                 720

Val Leu Ser Lys Ile Ala Leu Gly Asp Ser Ile Met Lys Pro Val Ser
                725                 730                 735

Pro
```

<210> 21
<211> 718
<212> PRT
<213> Mus musculus

<400> 21

```
Asp Ser Ser Met Val Ala Pro Leu Cys Pro Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Val
            20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Asp Asn Cys Gly Asp Thr Ser
        35                  40                  45

Gly Trp Thr Thr Ile Phe Gly Thr Val His Gly Asn Val Asn Lys Val
    50                  55                  60

Thr Leu Thr Gln Glu Cys Phe Leu Ser Gln Tyr Pro Gln His Cys Tyr
65                  70                  75                  80

Cys Arg Glu Asn Glu Leu Glu Cys Val Lys Ala Asp Leu Lys Ala Val
                85                  90                  95

Pro Lys Val Ser Ser Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Arg Leu Pro Val Lys Val Phe Ser Arg Tyr Thr Glu Leu Arg
        115                 120                 125

Lys Ile Tyr Leu Gln His Asn Cys Ile Thr His Ile Ser Arg Arg Ala
    130                 135                 140

Phe Leu Gly Leu His Asn Leu Gln Ile Leu Tyr Leu Ser His Asn Cys
145                 150                 155                 160

Ile Thr Ser Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Ala
                165                 170                 175

Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Lys Ser
```

180 185 190

Phe Met Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Gly Asn Arg
195 200 205

Leu Glu Ala Leu Pro Glu Thr Leu Cys Ala Gln Met Pro Gln Leu Asn
210 215 220

Trp Val Asp Leu Ala Asn Asn Gly Ile Lys Tyr Ile Thr Asn Ser Thr
225 230 235 240

Phe Leu Thr Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
245 250 255

Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
260 265 270

Glu Leu Asp Leu Ser Ser Asn Met Ile Thr Lys Leu Pro Val His Leu
275 280 285

Phe Ser Asp Leu His Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
290 295 300

Leu Leu Tyr Val His Lys Asn Gln Phe Gly Ser Leu Lys Gln Leu Gln
305 310 315 320

Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Ser Thr Gly Met
325 330 335

Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Leu Lys Thr Phe Arg
340 345 350

Tyr Cys Ser Tyr Val Pro His Val Arg Ile Cys Met Pro Ser Thr Asp
355 360 365

Gly Ile Ser Ser Ser Glu Asp Leu Leu Ala Asn Gly Ile Leu Arg Val
370 375 380

Ser Val Trp Val Ile Ala Phe Ile Thr Cys Val Gly Asn Phe Leu Val
385 390 395 400

Ile Ala Val Arg Ser Leu Ile Lys Ala Glu Asn Thr Thr His Ala Met
405 410 415

Ser Ile Lys Ile Leu Cys Cys Ala Asp Cys Leu Met Gly Val Tyr Leu
420 425 430

Phe Ser Val Gly Val Phe Asp Ile Lys Tyr Arg Gly Gln Tyr Gln Lys
435 440 445

Tyr Ala Leu Leu Trp Met Glu Ser Val Pro Cys Arg Leu Leu Gly Phe
450 455 460

Leu Ala Thr Leu Ser Thr Glu Val Ser Val Leu Leu Leu Thr Phe Leu
465 470 475 480

Thr Leu Glu Lys Phe Leu Val Ile Val Phe Pro Phe Ser Asn Leu Arg
485 490 495

Leu Gly Lys Arg Gln Thr Ala Val Ala Leu Ala Ser Ile Trp Val Val
500 505 510

```
Gly Phe Leu Ile Ala Ala Val Pro Phe Thr Arg Glu Asp Tyr Phe Gly
        515                 520                 525

Asn Phe Tyr Gly Lys Asn Gly Val Cys Phe Pro Leu His Tyr Asp Gln
    530                 535                 540

Ala Glu Asp Phe Gly Ser Arg Gly Tyr Ser Leu Gly Ile Phe Leu Gly
545                 550                 555                 560

Val Asn Leu Leu Ala Phe Leu Val Ile Val Ile Ser Tyr Val Thr Met
                565                 570                 575

Phe Cys Ser Ile His Lys Thr Ala Leu Gln Thr Ala Glu Val Arg Ser
            580                 585                 590

His Ile Gly Lys Glu Val Ala Val Ala Asn Arg Phe Phe Phe Ile Val
        595                 600                 605

Phe Ser Asp Ala Ile Cys Trp Ile Pro Val Phe Val Val Lys Ile Leu
    610                 615                 620

Ser Leu Leu Gln Val Glu Ile Pro Gly Thr Ile Thr Ser Trp Ile Val
625                 630                 635                 640

Val Phe Phe Leu Pro Val Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr
                645                 650                 655

Leu Thr Thr Ser Phe Phe Lys Asp Lys Leu Lys Gln Leu Leu His Lys
            660                 665                 670

His Arg Arg Lys Pro Ile Phe Lys Val Lys Lys Lys Ser Leu Ser Ala
        675                 680                 685

Ser Ile Val Trp Thr Asp Glu Ser Ser Leu Lys Leu Gly Val Leu Ser
    690                 695                 700

Lys Ile Ala Leu Gly Asp Ser Ile Met Lys Pro Val Ser Pro
705                 710                 715
```

<210> 22
<211> 1140
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (1)..(1140)

<400> 22

```
gac agc agt atg gtg gct cct ctg tgc ccc aaa ggg tat ttt ccc tgt    48
Asp Ser Ser Met Val Ala Pro Leu Cys Pro Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

ggg aat ctc acc aaa tgc ttg ccc cga gcc ttt cac tgc gat ggt gtg    96
Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Val
             20                  25                  30

gat gat tgc ggg aat ggt gcc gac gag gac aac tgt ggt gac act agt   144
Asp Asp Cys Gly Asn Gly Ala Asp Glu Asp Asn Cys Gly Asp Thr Ser
```

```
          35                  40                  45

gga tgg aca acc ata ttt ggc aca gtc cat ggg aat gtc aat aaa gtg   192
Gly Trp Thr Thr Ile Phe Gly Thr Val His Gly Asn Val Asn Lys Val
     50                  55                  60

aca ttg aca cag gag tgc ttt ctc agc cag tat cca cag cac tgt tac   240
Thr Leu Thr Gln Glu Cys Phe Leu Ser Gln Tyr Pro Gln His Cys Tyr
 65                  70                  75                  80

tgc aga gaa aat gaa ctg gaa tgt gta aag gct gac tta aaa gct gtg   288
Cys Arg Glu Asn Glu Leu Glu Cys Val Lys Ala Asp Leu Lys Ala Val
                 85                  90                  95

cca aag gtt tcc agc aac gta aca tta cta tct ctt aag aaa aac aaa   336
Pro Lys Val Ser Ser Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

atc cac aga ctt cca gtc aag gtc ttc agc aga tac aca gaa ctc aga   384
Ile His Arg Leu Pro Val Lys Val Phe Ser Arg Tyr Thr Glu Leu Arg
        115                 120                 125

aag ata tac ctt cag cac aac tgc atc aca cac atc tcc agg aga gca   432
Lys Ile Tyr Leu Gln His Asn Cys Ile Thr His Ile Ser Arg Arg Ala
    130                 135                 140

ttc ctt gga tta cat aat cta caa ata ctg tat ctc agc cat aac tgc   480
Phe Leu Gly Leu His Asn Leu Gln Ile Leu Tyr Leu Ser His Asn Cys
145                 150                 155                 160

att acc tct ctc agg cct ggg ata ttc aaa gac ttg cat cag ctt gct   528
Ile Thr Ser Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Ala
                165                 170                 175

tgg cta att tta gat gac aac ccg atc acc aga atc tca cag aag tcc   576
Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Lys Ser
            180                 185                 190

ttt atg ggg tta aac tcc ttg ttt ttc ttg tcc atg gtg ggt aac cgg   624
Phe Met Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Gly Asn Arg
        195                 200                 205

ctc gag gcc ctt cct gaa aca ttg tgt gct cag atg cct caa ctc aac   672
Leu Glu Ala Leu Pro Glu Thr Leu Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

tgg gtg gat ctg gca aac aat gga ata aag tac ata acg aac tcc acc   720
Trp Val Asp Leu Ala Asn Asn Gly Ile Lys Tyr Ile Thr Asn Ser Thr
225                 230                 235                 240

ttc cta acg tgc gac tcg ctc acg gtt ctg ttt ctg cct aga aat caa   768
Phe Leu Thr Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
                245                 250                 255

att ggt ttt gtt cca gag aag aca ttt tct tca tta aaa aat tta gga   816
Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
            260                 265                 270

gaa ctg gac ctg tct agc aat atg ata aca aaa ctc cca gtc cac ctt   864
Glu Leu Asp Leu Ser Ser Asn Met Ile Thr Lys Leu Pro Val His Leu
        275                 280                 285
```

```
ttc agc gac ctt cat ctt ctc cag aag ctg aac ctg tca tcc aac cct   912
Phe Ser Asp Leu His Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
    290                 295                 300

ctt ctg tat gtc cac aag aac cag ttt gga agt ctc aaa caa ctt cag   960
Leu Leu Tyr Val His Lys Asn Gln Phe Gly Ser Leu Lys Gln Leu Gln
305                 310                 315                 320

tct cta gac ctg gaa agg ata gag att cca aac ata agc aca gga atg   1008
Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Ser Thr Gly Met
                325                 330                 335

ttc cag cca atg aag aac ctt tct cac att tat ttg aaa acc ttt cga   1056
Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Leu Lys Thr Phe Arg
            340                 345                 350

tac tgc tcc tat gtc ccc cat gtc cga atc tgt atg ccg tcg act gat   1104
Tyr Cys Ser Tyr Val Pro His Val Arg Ile Cys Met Pro Ser Thr Asp
        355                 360                 365

ggt att tct tcg tct gag gac ctc ttg gct aac ggt                   1140
Gly Ile Ser Ser Ser Glu Asp Leu Leu Ala Asn Gly
    370                 375                 380
```

<210> 23
<211> 380
<212> PRT
<213> Mus musculus

<400> 23

```
Asp Ser Ser Met Val Ala Pro Leu Cys Pro Lys Gly Tyr Phe Pro Cys
  1               5                  10                  15

Gly Asn Leu Thr Lys Cys Leu Pro Arg Ala Phe His Cys Asp Gly Val
            20                  25                  30

Asp Asp Cys Gly Asn Gly Ala Asp Glu Asp Asn Cys Gly Asp Thr Ser
        35                  40                  45

Gly Trp Thr Thr Ile Phe Gly Thr Val His Gly Asn Val Asn Lys Val
    50                  55                  60

Thr Leu Thr Gln Glu Cys Phe Leu Ser Gln Tyr Pro Gln His Cys Tyr
65                  70                  75                  80

Cys Arg Glu Asn Glu Leu Glu Cys Val Lys Ala Asp Leu Lys Ala Val
                85                  90                  95

Pro Lys Val Ser Ser Asn Val Thr Leu Leu Ser Leu Lys Lys Asn Lys
            100                 105                 110

Ile His Arg Leu Pro Val Lys Val Phe Ser Arg Tyr Thr Glu Leu Arg
        115                 120                 125

Lys Ile Tyr Leu Gln His Asn Cys Ile Thr His Ile Ser Arg Arg Ala
    130                 135                 140

Phe Leu Gly Leu His Asn Leu Gln Ile Leu Tyr Leu Ser His Asn Cys
145                 150                 155                 160
```

```
Ile Thr Ser Leu Arg Pro Gly Ile Phe Lys Asp Leu His Gln Leu Ala
                165                 170                 175

Trp Leu Ile Leu Asp Asp Asn Pro Ile Thr Arg Ile Ser Gln Lys Ser
            180                 185                 190

Phe Met Gly Leu Asn Ser Leu Phe Phe Leu Ser Met Val Gly Asn Arg
        195                 200                 205

Leu Glu Ala Leu Pro Glu Thr Leu Cys Ala Gln Met Pro Gln Leu Asn
    210                 215                 220

Trp Val Asp Leu Ala Asn Asn Gly Ile Lys Tyr Ile Thr Asn Ser Thr
225                 230                 235                 240

Phe Leu Thr Cys Asp Ser Leu Thr Val Leu Phe Leu Pro Arg Asn Gln
                245                 250                 255

Ile Gly Phe Val Pro Glu Lys Thr Phe Ser Ser Leu Lys Asn Leu Gly
            260                 265                 270

Glu Leu Asp Leu Ser Ser Asn Met Ile Thr Lys Leu Pro Val His Leu
        275                 280                 285

Phe Ser Asp Leu His Leu Leu Gln Lys Leu Asn Leu Ser Ser Asn Pro
    290                 295                 300

Leu Leu Tyr Val His Lys Asn Gln Phe Gly Ser Leu Lys Gln Leu Gln
305                 310                 315                 320

Ser Leu Asp Leu Glu Arg Ile Glu Ile Pro Asn Ile Ser Thr Gly Met
                325                 330                 335

Phe Gln Pro Met Lys Asn Leu Ser His Ile Tyr Leu Lys Thr Phe Arg
            340                 345                 350

Tyr Cys Ser Tyr Val Pro His Val Arg Ile Cys Met Pro Ser Thr Asp
        355                 360                 365

Gly Ile Ser Ser Ser Glu Asp Leu Leu Ala Asn Gly
    370                 375                 380


<210> 24
<211> 757
<212> PRT
<213> Homo sapiens

<400> 24
Met Thr Ser Gly Ser Val Phe Phe Tyr Ile Leu Ile Phe Gly Lys Tyr
  1               5                  10                  15

Phe Ser His Gly Gly Gly Gln Asp Val Lys Cys Ser Leu Gly Tyr Phe
             20                  25                  30

Pro Cys Gly Asn Ile Thr Lys Cys Leu Pro Gln Leu Leu His Cys Asn
         35                  40                  45

Gly Val Asp Asp Cys Gly Asn Gln Ala Asp Glu Asp Asn Cys Gly Asp
     50                  55                  60
```

```
Asn Asn Gly Trp Ser Met Gln Phe Asp Lys Tyr Phe Ala Ser Tyr Tyr
 65                  70                  75                  80

Lys Met Thr Ser Gln Tyr Pro Phe Glu Ala Glu Thr Pro Glu Cys Leu
                 85                  90                  95

Val Gly Ser Val Pro Val Gln Cys Leu Cys Gln Gly Leu Glu Leu Asp
            100                 105                 110

Cys Asp Glu Thr Asn Leu Arg Ala Val Pro Ser Val Ser Ser Asn Val
        115                 120                 125

Thr Ala Met Ser Leu Gln Trp Asn Leu Ile Arg Lys Leu Pro Pro Asp
    130                 135                 140

Cys Phe Lys Asn Tyr His Asp Leu Gln Lys Leu Tyr Leu Gln Asn Asn
145                 150                 155                 160

Lys Ile Thr Ser Ile Ser Ile Tyr Ala Phe Arg Gly Leu Asn Ser Leu
                165                 170                 175

Thr Lys Leu Tyr Leu Ser His Asn Arg Ile Thr Phe Leu Lys Pro Gly
            180                 185                 190

Val Phe Glu Asp Leu His Arg Leu Glu Trp Leu Ile Ile Glu Asp Asn
        195                 200                 205

His Leu Ser Arg Ile Ser Pro Pro Thr Phe Tyr Gly Leu Asn Ser Leu
    210                 215                 220

Ile Leu Leu Val Leu Met Asn Asn Val Leu Thr Arg Leu Pro Asp Lys
225                 230                 235                 240

Pro Leu Cys Gln His Met Pro Arg Leu His Trp Leu Asp Leu Glu Gly
                245                 250                 255

Asn His Ile His Asn Leu Arg Asn Leu Thr Phe Ile Ser Cys Ser Asn
            260                 265                 270

Leu Thr Val Leu Val Met Arg Lys Asn Lys Ile Asn His Leu Asn Glu
        275                 280                 285

Asn Thr Phe Ala Pro Leu Gln Lys Leu Asp Glu Leu Asp Leu Gly Ser
    290                 295                 300

Asn Lys Ile Glu Asn Leu Pro Pro Leu Ile Phe Lys Asp Leu Lys Glu
305                 310                 315                 320

Leu Ser Gln Leu Asn Leu Ser Tyr Asn Pro Ile Gln Lys Ile Gln Ala
                325                 330                 335

Asn Gln Phe Asp Tyr Leu Val Lys Leu Lys Ser Leu Ser Leu Glu Gly
            340                 345                 350

Ile Glu Ile Ser Asn Ile Gln Gln Arg Met Phe Arg Pro Leu Met Asn
        355                 360                 365

Leu Ser His Ile Tyr Phe Lys Lys Phe Gln Tyr Cys Gly Tyr Ala Pro
    370                 375                 380
```

His Val Arg Ser Cys Lys Pro Asn Thr Asp Gly Ile Ser Ser Leu Glu
385 390 395 400

Asn Leu Leu Ala Ser Ile Ile Gln Arg Val Phe Val Trp Val Val Ser
405 410 415

Ala Val Thr Cys Phe Gly Asn Ile Phe Val Ile Cys Met Arg Pro Tyr
420 425 430

Ile Arg Ser Glu Asn Lys Leu Tyr Ala Met Ser Ile Ile Ser Leu Cys
435 440 445

Cys Ala Asp Cys Leu Met Gly Ile Tyr Leu Phe Val Ile Gly Gly Phe
450 455 460

Asp Leu Lys Phe Arg Gly Glu Tyr Asn Lys His Ala Gln Leu Trp Met
465 470 475 480

Glu Ser Thr His Cys Gln Leu Val Gly Ser Leu Ala Ile Leu Ser Thr
485 490 495

Glu Val Ser Val Leu Leu Leu Thr Phe Leu Thr Leu Glu Lys Tyr Ile
500 505 510

Cys Ile Val Tyr Pro Phe Arg Cys Val Arg Pro Gly Lys Cys Arg Thr
515 520 525

Ile Thr Val Leu Ile Leu Ile Trp Ile Thr Gly Phe Ile Val Ala Phe
530 535 540

Ile Pro Leu Ser Asn Lys Glu Phe Phe Lys Asn Tyr Tyr Gly Thr Asn
545 550 555 560

Gly Val Cys Phe Pro Leu His Ser Glu Asp Thr Glu Ser Ile Gly Ala
565 570 575

Gln Ile Tyr Ser Val Ala Ile Phe Leu Gly Ile Asn Leu Ala Ala Phe
580 585 590

Ile Ile Ile Val Phe Ser Tyr Gly Ser Met Phe Tyr Ser Val His Gln
595 600 605

Ser Ala Ile Thr Ala Thr Glu Ile Arg Asn Gln Val Lys Lys Glu Met
610 615 620

Ile Leu Ala Lys Arg Phe Phe Phe Ile Val Phe Thr Asp Ala Leu Cys
625 630 635 640

Trp Ile Pro Ile Phe Val Val Lys Phe Leu Ser Leu Leu Gln Val Glu
645 650 655

Ile Pro Gly Thr Ile Thr Ser Trp Val Val Ile Phe Ile Leu Pro Ile
660 665 670

Asn Ser Ala Leu Asn Pro Ile Leu Tyr Thr Leu Thr Thr Arg Pro Phe
675 680 685

Lys Glu Met Ile His Arg Phe Trp Tyr Asn Tyr Arg Gln Arg Lys Ser
690 695 700

Met Asp Ser Lys Gly Gln Lys Thr Tyr Ala Pro Ser Phe Ile Trp Val

705 710 715 720

Glu Met Trp Pro Leu Gln Glu Met Pro Pro Glu Leu Met Lys Pro Asp
725 730 735

Leu Phe Thr Tyr Pro Cys Glu Met Ser Leu Ile Ser Gln Ser Thr Arg
740 745 750

Leu Asn Ser Tyr Ser
755

<210> 25
<211> 11
<212> PRT
<213> Human immunodeficiency virus type 1

<400> 25
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
1 5 10

<210> 26
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: internalizing domain derived from HIV tat protein

<400> 26
Gly Gly Gly Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
1 5 10 15

<210> 27
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 27
tgccaaaaag gatattttcc ctgtgggaat ctta 34

<210> 28
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 28
ctaggaaact ggtttcatta tactgtctcc aagtgttatt ttgttca 47

<210> 29

<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: RACE primer

<400> 29
ccatcctaat acgactcact atagggc 27

<210> 30
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: RACE primer

<400> 30
attgtcatct agaattagcc aagttagctg at 32

<210> 31
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: RACE primer

<400> 31
aacaaggaat ttaatcccgt aaacaag 27

<210> 32
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: RACE primer

<400> 32
actcactata gggctcgagc ggc 23

<210> 33
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: RACE primer

<400> 33
atattccagg tctgagggtt gtgat 25

<210> 34
<211> 25

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: RACE primer

<400> 34
atattccagg tctgagggtt gtgat 25

<210> 35
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 35
ctgctttgga aatctttttg tca 23

<210> 36
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 36
ttttccaggt cgaatgttac tga 23

<210> 37
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 37
atgccttgct gtggatggag 20

<210> 38
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 38
acttcggtgg acagcatgg 19

<210> 39
<211> 20
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide probe

<400> 39
cgtgcagtgc cgcctcatgg 20

<210> 40
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 40
gtcgacggcg agccc 15

<210> 41
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 41
tctttgggac cttgtctgca a 21

<210> 42
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide probe

<400> 42
tgggccgcgt ctcctttgag ct 22